National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

UNIVERSITY OF ALBERTA

Development of a Miniaturized Peptide Sequencer with Capillary Electrophoresis Separation and Thermo-Optical Absorbance Detection of Derivatized Amino Acids

BY ©

Karen C. Waldron

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

DEPARTMENT OF CHEMISTRY
EDMONTON, ALBERTA
SPRING 1994

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-11403-1

Canada

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Karen C. Waldron

TITLE OF THESIS: Development of a Miniaturized Peptide Sequencer with Capillary Electrophoresis Separation and Thermo-Optical Absorbance Detection of Derivatized Amino Acids

DEGREE: Doctor of Philosophy

YEAR THIS DEGREE GRANTED: 1994

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Karen Waldron

8640-108 Street, Suite 202
Edmonton, Alberta T6E 4M4

Date Jan. 4, 1994.

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Development of a Miniaturized Peptide Sequencer with Capillary Electrophoresis Separation and Thermo-Optical Absorbance Detection of Derivatized Amino Acids** submitted by **Karen C. Waldron** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Dr. N.J. Dovichi

Gary Horlick
Dr. G. Horlick

George Kotovych
Dr. G. Kotovych

Dr. L. Li

Dr. L.B. Smillie

R. Aebersold
Dr. R. Aebersold

Date Jan. 4, 1994.

In loving memory of my mother

Abstract

Analysis of ever decreasing amounts of sample requires ever increasing sophistication in analytical techniques. This thesis describes the use of capillary electrophoresis (CE) separation and laser-based thermo-optical detection for the analysis of minute amounts of biologically important samples: purine and pyrimidine bases, nucleosides, DNA, and derivatized amino acids specifically relevant to peptide and protein sequencing. This thesis also describes the development of a miniaturized peptide sequencing instrument designed to match the volume requirements of CE and permit high sensitivity determination of the derivatized amino acid products from sequencing.

Peptide (and protein) sequencing is the process of determining primary amino acid sequence. Knowledge of primary sequence is important for revealing evolutionary history, identifying abnormal function and providing access to genes. Researchers are always trying to find ways of improving the sensitivity of peptide sequencing, limited now to about one picomole of peptide. Traditional methods of identifying the sequencing products, phenylthiohydantoin (PTH) amino acids, revolve around chromatography. Commercial sequencing instruments use reversed-phase high performance liquid chromatography (HPLC) with ultraviolet (UV) absorbance but the detection ability of this technique is just under one picomole. This thesis explores the use of CE as an alternative separation scheme and laser-based thermo-optical absorbance as an alternative detection scheme for PTH amino acids.

CE provides high speed, high efficiency separations compared to HPLC; analysis times are 5 times faster and peak efficiencies are 10 to 100 times higher. Laser-based thermo-optical detection is an absorbance based method of detection independent of optical path length so high sensitivity detection in capillary tubes is possible. Limits of detection are 1,000 times better with thermo-optical absorbance compared to conventional UV detectors used in HPLC. In this thesis, CE with thermo-optical detection of PTH amino acids from peptide sequencing show superior separation and detection compared to state-of-the-art HPLC.

Interface of the CE system directly to a commercial peptide sequencer to improve sequencing sensitivity is not possible because the sample volume requirements differ by four to five orders of magnitude. The design and characterization of a highly miniaturized peptide sequencer are presented in this thesis, with the intention of matching the volume requirements more closely to CE analysis. The development of micro-scale analytical techniques, like those described in this thesis, will provide new avenues to study samples that are difficult to isolate or purify in large amounts.

Acknowledgments

I would like to thank all the people who have contributed to the research I've done in the past five years and to the preparation of this thesis—contributions such as chemical and technical advice, and moral support. Many people in and outside of Edmonton, friends and colleagues both, have helped me reach this final stage in my degree. I would like to express my gratitude to all these people, but I only have room to mention a few by name.

Above all, I would like to thank Dr. Norman Dovichi, my supervisor. Norm's expertise in science and research is matched only by his enthusiasm, qualities I hope to gain from having worked with him. He has the patience of Job and the hands of a sturgeon. Mostly, Norm's respect and encouragement have been generous towards building my confidence to finish this thesis.

The members of my Ph.D. committee, Dr. Gary Horlick, Dr. George Kotovych, Dr. Liang Li and Dr. Larry Smillie, have been very helpful with their comments and encouragement; I wish to thank them all. Dr. Smillie helped me understand the biochemistry that I constantly struggled with and Dr. Horlick did his best to correct the chemist's mutation of the Queen's english.

Each of the past and present members of The Northern Lights Laser Lab are due a load of thanks for their help, understanding and tolerance of my endless questions and boisterous nature. Thanks guys. I would also like to thank Michael Carpenter from the biochemistry department at U of A and Ted Bures from the Biomedical Research Centre at UBC. These two gentlemen spent several hours helping me understand the nuances of gas-phase protein sequencing and provided me with samples and sequence results to use in my research.

I would like to thank all the support staff of the chemistry department at U of A who helped me build things, fix things, photocopy things, buy things and make sure I got paid. I am especially in debt to Hubert Hoffmann's team in the machine shop and Frank Dreissigacker's team in the electronics shop for their technical help.

Finally, I thank with all my heart, my family. My father, Keith Cameron, deserves thanks for his patience and endless explaining to his friends as to why his 30-year-old daughter is *still* in school. Most of all, my husband David deserves a million thanks for providing love, moral support and encouragement throughout my degree program. Even though David could not pronounce fifty percent of the words in this thesis, he tirelessly read through it attempting to correct my appalling grammar. Thanks honey.

Table of Contents

1 Introduction 1
  1.1 Introduction 2
  1.2 Capillary Electrophoresis 3
  1.3 Thermo-Optical Absorbance Detection 14
  1.4 Peptide and Protein Sequence Analysis 27
  1.5 References 39
  Appendix A: L-α-amino acid structures 43
  Appendix B: N-terminal derivatizing reagents 44

2 Capillary Electrophoresis and Thermo-Optical Absorbance Detection of Amino Acid Derivatives 45
  2.1 Introduction 46
  2.2 Dimethylaminoazobenzene Thiohydantoin (DABTH) Amino Acids 47
    2.2.1 Experimental 48
    2.2.2 Results and Discussion 49
  2.3 Phenylthiohydantoin (PTH) Amino Acids 53
    2.3.1 Experimental 55
    2.3.2 Results and Discussion 56
  2.4 Conclusions 68
  2.5 References 71

3 Capillary Electrophoresis and Thermo-Optical Absorbance Detection of Products from Manual Peptide Sequencing 73
  3.1 Introduction 74
  3.2 Experimental 77
  3.3 Results and Discussion 79
  3.4 Conclusions 102
  3.5 References 103
  Appendix C: Normalization of migration times 107

4 Development of a Highly Miniaturized Peptide Sequencer 109
  4.1 Introduction 110
  4.2 Experimental 115
  4.3 Results and Discussion 125

4.4 Conclusions 141
4.5 References 143

5 Capillary Electrophoresis and Thermo-Optical Absorbance Detection of Bases, Nucleosides and DNA 147
5.1 Introduction 148
5.2 Experimental 152
5.3 Results and Discussion 155
5.4 Conclusions 170
5.5 References 171

6 Conclusions and Future Directions 174
6.1 Discussion 175
6.2 References 181

List of Figures

1.1 Number of publication titles containing the words capillary and electrophoresis for years 1980 to 1992 4

1.2 Capillary zone electrophoresis apparatus 5

1.3 Electroosmotic versus parabolic flow 7

1.4 Micellar electrokinetic capillary chromatography using SDS 13

1.5 The crossed beam thermal lens 20

1.6 Experimental setup for capillary thermo-optical absorbance detector 24

1.7 Detailed thermo-optical absorbance detection 25

1.8 Photograph of probe beam profile after diffraction and refraction by sample in the capillary tube 26

1.9 Scheme for sequencing a peptide 30

1.10 The Edman degradation 32

2.1 Reaction of dimethylaminoazobenzene isothiocyanate (DABITC) with an amino acid 49

2.2 Spectrophotometric study to determine the pKa of DABTH-tyrosine: (A) absorbance spectrum of DABTH-tyrosine at pH 1.75, (B) absorbance spectrum of DABTH-tyrosine at pH 6.88, (C) plot of the shift in wavelength with pH for the absorbance of DABTH-tyrosine 51

2.3 Capillary electropherogram of DABTH amino acids 52

2.4 Block diagram of experimental setup 56

2.5 Reaction of phenylisothiocyanate (PITC) with an amino acid 57

2.6 High-speed MECC electropherogram of PTH amino acids, 60 fmol each 59

2.7 High resolution MECC electropherogram of PTH amino acids 60

2.8 Absorbance spectrum of PTH glycine 61

2.9 Plot of time resolved thermo-optical signal generated by a pulsed excimer waveguide laser operating at 1,000 pulses per second 62

2.10 Calibration curve for PTH-glycine 63

2.11 Plot of linear dynamic range for PTH glycine 67

3.1 Amino acid sequences of insulin chain B and SP-5, splenopentin 80

3.2 By-products formed during the coupling step of Edman degradation 82

3.3 Effect of SDS concentration on migration time of the PTH amino acids 84

3.4 Effect of pH on the migration time of PTH amino acids 85

3.5 Capillary electropherogram of 60 fmol PTH amino acid standard 87

3.6 Electropherogram of twenty PTH amino acids in organic/aqueous buffer 89

3.7 Electropherograms for the sequence analysis of SP-5 peptide 92

3.8A Electropherograms for the sequence analysis of insulin chain B, cycles 1 to 5 95

3.8B Electropherograms for the sequence analysis of insulin chain B, cycles 6 to 10........ 96

3.8C Electropherograms for the sequence analysis of insulin chain B, cycles 11 to 15 97

3.8D Electropherograms for the sequence analysis of insulin chain B, cycles 16 to 20 98

3.8E Electropherograms for the sequence analysis of insulin chain B, cycles 21 to 25 99

3.8F Electropherograms for the sequence analysis of insulin chain B, cycles 26, 27 and 29........ 100

4.1 Highly miniaturized peptide sequencer schematic 116

4.2 Diagram of reaction chamber........ 117

4.3 Thermoelectric modules for temperature control during coupling and cleavage reactions........ 118

4.4 Configuration of reagent and solvent delivery vials........ 121

4.5 Valco multiposition distribution valve........ 122

4.6 Voltage applied to thermoelectric modules versus temperature inside reaction chamber 126

4.7 Graph of temperature versus time after voltage is applied 127

4.8 Electropherograms showing the miniaturized peptide sequencer analysis of 50 pmol insulin chain B, (A) standard, (B) first cycle 133

4.9 Chromatograms showing the ABI Model 470A gas-phase sequencer analysis of insulin chain B 140

5.1 Schematic diagram of the primary structure of DNA 150

5.2 Structures of the bases and nucleosides 153

5.3 Double-stranded DNA with 3 base pairs shown 154

5.4 Electropherogram of a mixture of nine bases and nucleosides........ 157

5.5 Absorbance spectrum of adenine........ 158

5.6 Calibration curve for adenine........ 159

5.7 Plot of linear dynamic range of adenine 162

5.8 Slab gel electropherograms of molecular size markers used to size double-stranded DNA 164

5.9 Electropherogram of the 100-bp ladder 167

5.10 Electropherogram of the *Hae III* digest of ΦX174 RF DNA 168

6.1 Attomole peptide sequencer schematic........ 179

6.2 Schematic of a miniaturized peptide sequencer with mass spectral identification of amino acid derivatives.. 181

List of Tables

1.1 Methods of fragmentation for large polypeptides 37

2.1 Detection limits for PTH-amino acids 66

3.1 Quantitation of the sequence analysis of SP-5 peptide 90

3.2 Quantitation of the sequence analysis of insulin chain B 94

4.1 Length, size, and type of tubing used to deliver degradation reagents 120

4.2 Sequencer program 124

4.3 Comparison of reagent/solvent amounts for miniaturized sequencer versus other gas-phase sequencers 138

5.1 Calculation of detection limits for bases and nucleosides 161

Symbols and Abbreviations

The Common Amino Acids:

A	Ala	alanine
R	Arg	arginine
N	Asn	asparagine
D	Asp	aspartic acid
C	Cys	cysteine
E	Glu	glutamic acid
Q	Gln	glutamine
G	Gly	glycine
H	His	histidine
I	Ile	isoleucine
L	Leu	leucine
K	Lys	lysine
M	Met	methionine
F	Phe	phenylalanine
P	Pro	proline
S	Ser	serine
T	Thr	threonine
W	Trp	tryptophan
Y	Tyr	tyrosine
V	Val	valine

ABI	Applied Biosystems Inc.
ATZ	2-Anilino-5-Thiazolinone
BAMPITC	Boc-aminomethyl-phenyliosthiocyanate
BNPS-Skatole	2-(2-Nitrophenyl sulphenyl)-3-methyl-3'-bromoindolenine
CNBr	Cyanogen Bromide
CE	Capillary electrophoresis
CGE	Capillary gel electophoresis
CMC	Critical micelle concentration
CZE	Capillary zone electrophoresis
DABITC	4-*N,N*-Dimethylaminoazo-benzene-4'-isothiocyanate
DABSYL	4,4-Dimethylaminoazobenzene-4'-sulfonyl
DABTH	Dimethylaminoazobenzene thiohydantoin
DANSYL	5-Dimethylamino naphthalene-1-sulfonyl
DITC	Diphenyl isothiocyanate
DMPTU	N,N-Dimethyl-N'-phenylthiourea
DNA	Deoxyribonucleic acid
DNSAPITC	Dimethylaminoaphthylsulfonylamino-phenylisothiocyanate
DPTU	Diphenylthiourea
DTT	Dithiothreatol

EA	Ethyl acetate
FITC	Fluoresceinisothiocyanate
HPLC	High performance liquid chromatograpy
ID	Inner diameter
IEF	Isoelectric focusing
ITP	Isotachophoresis
MECC	Micellar electrokinetic capillary chromatography
MeCN	Acetonitrile
MITC	Methylisothiocyanate
MS	Mass spectrometry
OD	Outer diameter
PAGE	Polyacrylamide gel electrophoresis
PETAPITC	3-[4'(ethylene-*N,N,N*-trimethylamino)phenyl]-2-isothiocyanate
PCR	Polymerase chain reaction
Polybrene®	Hexadimethrine Bromide
PTC	Phenylthiocarbamyl
PTH	Phenylthiohydantoin
PTFE	Polytetrafluoroethylene
PTU	Phenylthiourea
PVDF	Polyvinylidene difluoride
RNA	Ribonucleic acid
RY	Repetitive yield
SDS	Sodium dodecyl sulfate
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TMA	Trimethylamine
TFA	Trifluoroacetic acid
UV	Ultraviolet

CHAPTER 1

INTRODUCTION

1.1 Introduction

"If it were possible to obtain sequence information on [one to ten femtomoles of protein] the impact on biology and medicine would be tremendous, and we would see the establishment of new fields of research" [1]. Peptides and proteins constitute over eighty percent of the organic molecules in the human body and play an enormous variety of roles, from transport and storage of small molecules to immune response. Determination of the primary structure of peptides and proteins is a fundamental part of understanding protein structure and function. As Lloyd Smith implies in the quote above, if the primary sequence of very small amounts of protein could be determined then previously inaccessible biomolecules could be studied. It is the development of micro-scale analytical techniques, like those described in this thesis, that will enable sequence analysis of femtomole amounts of peptides and proteins.

The micro-scale analytical techniques described here are capillary electrophoresis, laser-based detection and polypeptide microsequencing. State-of-the-art polypeptide sequencing involves derivatization and cleavage of the terminal amino acid from a peptide followed by identification of the cleaved amino acid. Determination of the derivatized amino acids is typically done by high performance liquid chromatography (HPLC) with ultraviolet (UV) absorbance detection; however, identification of femtomole quantities is difficult. Capillary electrophoresis (CE) is a relatively new, high-efficiency separation technique, complementary to high performance liquid chromatography (HPLC). CE is particularly well suited to analysis of less than one picomole of sample and is compatible with laser-based detection schemes.

Laser-based thermo-optical absorbance detection in capillaries offers better sensitivity than conventional absorbance detection. It is well suited to determination of the derivatized amino acid products from sequencing. Initial work on developing a separation and detection scheme for dimethylaminoazobenzene thiohydantoin derivatized amino acids is presented in the first part of Chapter 2. Building a detector to identify phenylthiohydantoin (PTH) amino acids was made possible by the recent commercial availability of an inexpensive UV excimer laser. This work is discussed in the second part of Chapter 2. In Chapter 3, the feasibility of CE/laser-based detection of PTH amino acids is applied to manual peptide sequence analysis. Chapter 4 describes the design and characterization of a micro-scale sequencer compatible with the volume requirements of the CE/laser-based analysis system. In Chapter 5, identification of a few other important biomolecules is explored using the novel UV laser and capillary electrophoresis.

The work presented in this thesis draws together the speed and efficiency of capillary electrophoresis, the high sensitivity of thermo-optical absorbance detection and the principles of polypeptide sequencing to advance the field of microsequencing and hopefully help to establish new fields of research in the biosciences.

1.2 Capillary Electrophoresis

Electrophoresis has been a tool of the biological sciences since the 1930s [2]. However, the past fifteen years have seen the rapid growth of capillary electrophoresis, a technique first described by Hjerten in 1967 [3]. Figure 1.1 demonstrates the rate of publications cited in Chemical Abstracts by searching for 'capillary' and 'electrophoresis' appearing together. Between January 1990 and January 1992 alone there were almost eighty reviews covering capillary electrophoresis [4].

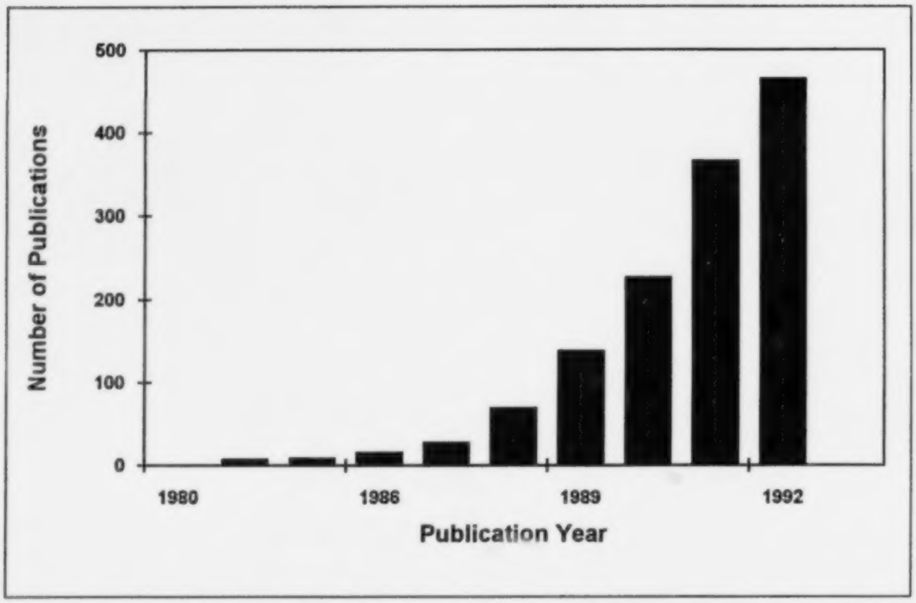


Figure 1.1 Number of publication titles containing the words ‘capillary’ and ‘electrophoresis’ for years 1980 to 1992.

The general principle of electrophoresis is the separation of charged species in an electrolytic medium under the influence of an externally applied electric field. Joule (resistive) heating generated by the applied field results in temperature gradients that degrade separation performance. Subsequent density gradients and convection also degrade the separation. Electrophoresis performed in capillaries that have a high surface-to-volume ratio (IDs from 2 μm to 250 μm) reduces Joule heating because of good heat transport. Better heat dissipation through the capillary walls permits use of higher electric fields leading to improved separations. The small inner diammeter of capillary tubes minimizes convection thereby reducing band broadening. Also, the ultrasmall volumes associated with capillary electrophoresis lend themselves well to determination of very small sample sizes in very small detection volumes.

The onslaught of capillary electrophoresis (CE) has brought us several similar modes: capillary zone electrophoresis (CZE), micellar electrokinetic capillary chromatography (MECC), capillary gel electrophoresis (CGE), capillary isotachophoresis (ITP) and capillary isoelectric focusing (IEF). While each type of CE has its own special features, this work will focus on CZE in order to present some theoretical concepts, and on MECC, which is the mode of electrophoresis employed for most of the experiments.

Capillary Zone Electrophoresis

Capillary zone electrophoresis (CZE) is an open tubular method for separating species based on their charge-to-mass ratio. Each end of the buffer-filled capillary is immersed in a vial of buffer with a high voltage electrode at the injection end and a grounded electrode at the detection end. A Plexiglas safety box is interlocked to the high voltage switch to protect the user during operation. The capillary tubes are almost exclusively made of fused silica. Current (i) is measured in the return circuit. A schematic of the basic CZE apparatus is shown in Figure 1.2.

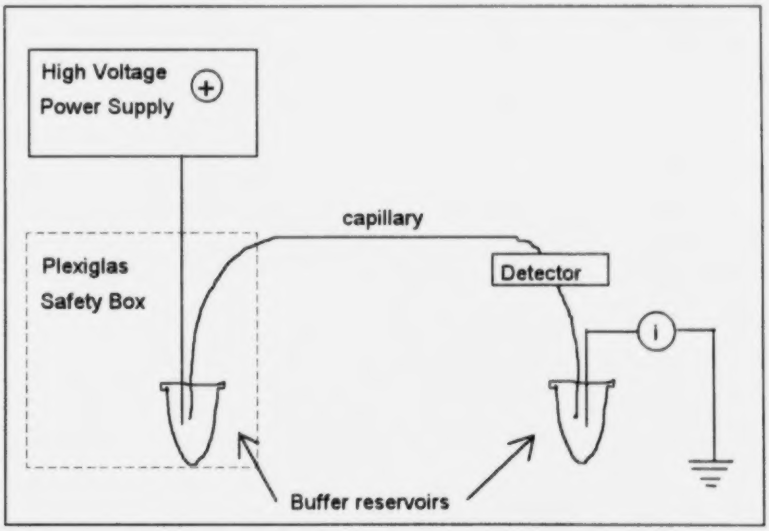


Figure 1.2 Capillary zone electrophoresis apparatus.

Besides electrophoresis, there is another phenomenon that contributes to the movement of species in CZE: electroosmosis. It is prudent to discuss the two processes together because the equations governing CZE separation contain terms from both electrophoresis and electroosmosis.

Briefly, electroosmosis is the bulk flow of solvent (buffer) in a capillary tube when an electric field is applied from one end to the other. It does not cause separation, but simply moves species along the capillary. This phenomenon occurs because a surface that is in contact with an aqueous solvent becomes charged. Oppositely charged solvent molecules are attracted to the surface and an electric double layer is formed. In fused silica tubing, the interior surface becomes net negatively charged because of silanol groups. Cations from the aqueous solvent line up along this surface to maintain electroneutrality and a electric double layer is formed. When an electric field is applied, the cationic solvent molecules are strongly attracted to the negative electrode and migrate toward it. This migration causes the rest of the aqueous solvent to be dragged toward the anode—electroosmosis. Several authors have dealt with this phenomenon in detail [2, 3, 5-8].

Figure 1.3 presents a simple diagram of electroosmotic flow compared to parabolic flow from a pressure pumped system such as liquid chromatography. In pressure driven flow, solvent at the tube centre moves with a maximum velocity based on pressure, tube radius and solvent viscosity. At the walls, velocity is zero. As the flow passes through the tube, viscous drag at the walls leads to the evolution of a parabolic flow profile. In voltage driven flow, solvent at the tube centre has zero velocity at time t_1 because it is effectively neutral in charge. The solvent cations in the electric double layer are attracted to the cathode and move with maximum velocity based on applied voltage, solvent viscosity and diffusion coefficient. At time t_2, the bulk solution begins to catch up with the cationic layer and the net effect is a flat flow profile, sometimes called plug flow.

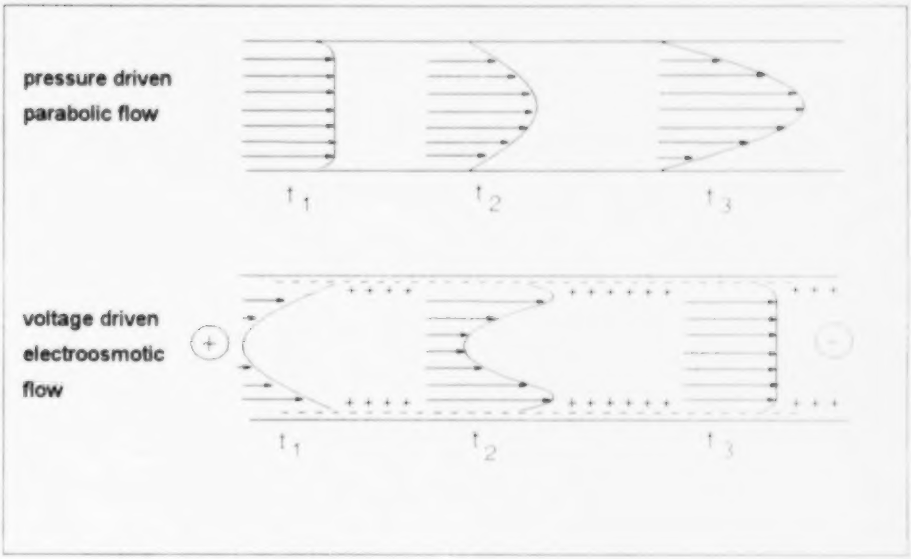


Figure 1.3 Electroosmotic versus parabolic flow [9].

Electroosmotic flow (EOF) is the velocity of the bulk solvent, v_{eo}. EOF is a function of the electroosmotic coefficient of mobility, μ_{eo}, and electric field strength, E, as defined below:

$$v_{eo} = \mu_{eo} \mathrm{E} = \frac{\varepsilon\zeta}{4\pi\eta} \mathrm{E} \qquad (1.1)$$

The electroosmotic coefficient of mobility, μ_{eo}, depends on pH and solvent composition where ε is the dielectric strength, η is the viscosity and ζ is the zeta potential measured at the plane of shear close to the solvent-surface interface. For the experiments described in Chapter 2, a typical EOF at pH 7 is about 1.4 mm/s. Zeta potentials are on the order of -10 to -150 mV [8].

As mentioned already, movement of ionic species in CZE is a function of both electrophoresis and electroosmosis. In the simplest sense, an ionic solute travels from the high voltage end to the detection end of the capillary with a given electrophoretic velocity, v_{ep}. Velocity depends on the electrophoretic coefficient of mobility of the solute, μ_{ep}, and the electric field strength, E, such that:

$$v_{ep} = \mu_{ep}\,E \tag{1.2}$$

where the unique electrophoretic mobilities of different species allows separation in CZE. An ion has a given μ_{ep} which contributes to its velocity of migration. Combined with the electroosmotic flow arising from the composition of the buffer, the linear velocity of charged species in CZE is defined as:

$$v = v_{ep} + v_{eo} = (\mu_{ep} + \mu_{eo})\,E = \left(\mu_{ep} + \mu_{eo}\right)\frac{V}{L_t} \tag{1.3}$$

where electric field strength, E, equals applied voltage, V, over total length of capillary, L_T. Even though an anionic solute is attracted to the positive electrode, the EOF is generally much stronger and all charged species are swept toward the negative electrode: $v_{eo} >> v_{ep}$. The migration time, t_m, for a species is simply distance traveled to the detector, L_d, divided by the velocity of travel, v. The result is that migration time in CZE is proportional to capillary length squared and inversely proportional to applied voltage. Faster separations are achieved in short capillaries run at high voltages:

$$t_m = \frac{L_d}{v} = \frac{L_d^{\,2}}{(\mu_{ep} + \mu_{eo})V} \tag{1.4}$$

Contrary to chromatographic theory, longer columns do not increase efficiency in CZE. In liquid chromatography, for example, several factors contribute to column bandbroadening such as longitudinal diffusion, non-uniform parabolic flow profile of velocities across the tube (trans-column resistance to mass transfer) and resistance to mass transfer in the stationary phase. In electroosmotically driven systems, the force is uniformly distributed across the entire length of the tube; there is no pressure drop available to create parabolic flow. Combined with the absence of a stationary phase, the major contribution to bandbroadening in CZE is longitudinal diffusion. Bandbroadening or peak variance, σ^2, can be described by the Stokes-Einstein equation:

$$\sigma^2 = 2\,D\,t_m \qquad (1.5)$$

where D is the diffusion coefficient of sample in the mobile phase measured in cm^2/s and t_m is the time that the diffusion was allowed to occur. A means of quantitatively describing the tendency of a particular chromatography column to produce bandbroadening is the plate number, N, also called the efficiency. Plate height, $\overline{H}$, or height equivalent to a theoretical plate (HETP) is the distance travelled by a solute, L_d, divided by the number of theoretical plates:

$$\overline{H} = \frac{L_d}{N} \qquad (1.6)$$

In the rate theory of column bandbroadening, plate height is defined as the variance of the band or peak per unit distance x_i along the column through which the band centre has migrated. When the centre of the band has reached the detector then $x_i = L_d$:

$$\overline{H} = \frac{\sigma^2}{L_d} \qquad (1.7)$$

Combining Equations 1.6 and 1.7, then substituting in Equation 1.5 gives the following expression for the number of theoretical plates or column efficiency in CZE:

$$N = \frac{L_d^{\,2}}{\sigma^2} = \frac{L_d^{\,2}}{2\,D\,t_m} = \frac{L_d^{\,2}}{2D\left(\dfrac{L_d^{\,2}}{(\mu_{ep} + \mu_{eo})V}\right)} = \frac{(\mu_{ep} + \mu_{eo})V}{2D} \qquad (1.8)$$

Equation 1.8 shows that column efficiency is a function of the electrophoretic and electroosmotic coefficients of mobility, the applied voltage and the solute diffusion coefficient, but not the column length. For this reason, high efficiency separations can be achieved in CZE by using high voltages, and fast separations can be achieved by using short capillary tubes (see Equation 1.4). The main reasons people use long capillaries ($>$20 cm) are that current and Joule heating are lower, and detectors are easier to install the further they are from the injection end. An easy method for calculating the plate number can be derived from the first part of Equation 1.8:

$$N = \frac{L_d^{\ 2}}{\sigma^2} \equiv \frac{t_m^{\ 2}}{\sigma_t^{\ 2}} \qquad (1.9)$$

Since the width at half-height, $W_{0.5}$, of a Gaussian peak is defined as $(8\ ln2)^{0.5}\times\sigma_t$, Equation 1.10 provides a quick, routine method of calculating plate number where t_m is migration time. For CZE, efficiencies are on the order of 10^5 to 10^6 theoretical plates.

$$N = 5.54\left(\frac{t_m}{W_{0.5}}\right)^2 \qquad (1.10)$$

Another important concept in separation theory is resolution. Resolution is the separation between peak centres in terms of the average baseline width of the peaks. So, for peaks i and j the resolution, R, is typically expressed as:

$$R = \frac{t_{m,j} - t_{m,i}}{\frac{1}{2}\left(W_i + W_j\right)} \qquad (1.11)$$

where $t_{m,j}$ and $t_{m,i}$ are the migration times and W_i and W_j are the baseline peaks widths. Since CZE peaks have migration times and not retention times as in conventional chromatography, it is also possible to represent resolution in terms of mobility. The electroosmotic coeffieicent of mobility is the same for both the i_{th} and j_{th} peak, and their peak widths will be approximately the same. The numerator in Equation 1.11 therefore reduces to $W_{i,j}$ which is equal to 4σ. From the first part of Equation 1.8 $\sigma=L/(N)^{0.5}$, therefore resolution can be expressed as the relative difference in mobilities of species i and j by Equation 1.12:

$$R = \frac{1}{4}\left(\frac{\Delta\mu_{ep}}{\overline{\mu}_{ep} + \mu_{eo}}\right)\sqrt{N} \tag{1.12}$$

where $\Delta\mu_{ep}$ equals $(\mu_{ep,j} - \mu_{ep,i})$ and $\overline{\mu}_{ep}$ is the average electrophoretic mobility [10]. If the expression for theoretical plates (Equation 1.8) is substituted into Equation 1.12, then the voltage has to be quadrupled just to double the resolution and therefore control of species' mobility is the key to improving resolution, as shown in the following equation:

$$R = \frac{1}{4\sqrt{2}}\left(\frac{\Delta\mu_{ep}}{(\overline{\mu}_{ep} + \mu_{eo})^{1/2}}\right)\frac{\sqrt{V}}{\sqrt{D}} \tag{1.13}$$

Control of mobility in CE can be effected by modifying the capillary inner surface or by modifying the buffer. The following section discusses one of the more common methods for changing the buffer to allow separation of neutral species.

Micellar Electrokinetic Capillary Chromatography

Equations 1.1 through 1.13 apply to charged species, in that a neutral species has zero electrophoretic mobility. To separate neutrals and molecules with very similar charge-to-mass ratios, addition of surfactant is helpful [11]. Above the critical micelle concentration (CMC), surfactant molecules aggregate to form spherically shaped micelles having a hydrophobic interior and charged exterior. As mentioned above, it is incorrect to talk about retention time in CZE because nothing is retained. In micellar electrokinetic capillary chromatography (MECC) though, micelles formed from the surfactant molecules provide a pseudo-phase into which analyte can partition, providing a means of retention. MECC is essentially voltage-pumped chromatography, in contrast to conventional liquid-pumped chromatography. Since chromatography requires two phases between which analyte is distributed, one can consider the buffer to be the 'mobile' phase and micelles the pseudo 'stationary' phase. Therefore, we can define

n_{mc} as the number of analyte molecules incorporated into the micelle and n_{aq} as the number in the aqueous phase. The migration velocity of the analyte can be expressed [12] in chromatographic terms as:

$$v = \left(\frac{n_{aq}}{n_{aq} + n_{mc}} \right) v_{eo} + \left(\frac{n_{mc}}{n_{aq} + n_{mc}} \right) v_{mc} \qquad (1.14)$$

where v_{mc} is the mobility of the analyte in the micelle. The micelles are essentially a stationary phase pumped through the capillary by electroosmotic flow. The capacity factor, k', is the ratio of the number of analyte molecules *inside* the micelle to those *outside* and can be expressed as:

$$k' = \frac{n_{mc}}{n_{aq}}. \qquad (1.15)$$

The migration velocity can also be expressed as length over time, by analogy to Equation 1.4, therefore:

$$v = \frac{L_d}{t_r} \quad \text{and} \quad v_{eo} = \frac{L_d}{t_0} \quad \text{and} \quad v_{mc} = \frac{L_d}{t_{mc}} \qquad (1.16)$$

where t_r is the observed retention time, t_0 is the retention time of the aqueous phase or unretained compound measured as EOF and t_{mc} is the retention time of the micelle. Combining Equations 1.14 through 1.16 gives an expression for the retention time of a neutral solute in MECC:

$$t_r = \frac{1 + k'}{1 + \left(\frac{t_0}{t_{mc}} \right) k'} t_0 \qquad (1.17)$$

Alternatively, this equation can be rearranged to calculate k' after measuring t_r, t_0 and t_{mc} empirically. The actual velocity of a neutral solute in MECC depends on electroosmosis too, and therefore, we can substitute the relationship between micellar and aqueous phase for μ_{ep} in Equation 1.3 to get:

$$v = \left(\mu_{eo} + \frac{k'}{1 + k'} \mu_{ep,mc} \right) E \qquad (1.18)$$

where $\mu_{ep,mc}$ is the electrophoretic coefficient of mobility of the micelle.

In the same manner as CZE, all species in MECC migrate from the positive electrode (anode) to the grounded or negative electrode—this is true even for anionic micelles. Sodium dodecyl sulfate (SDS) is perhaps the most popular surfactant for MECC and forms micelles when present above 8 mM. SDS forms large (radius ~3 nm) negatively charged micelles that are dragged along with the electroosmotic flow and elute at the detection end of the capillary. Figure 1.4 is a representation of MECC using SDS and shows the partitioning of neutral solutes in the micelle. Micelles have a substantial electrophoretic coefficient of mobility, directed toward the positive electrode. However, the electroosmotic flow (EOF) of the bulk solvent (buffer) is much stronger than the electrophoretic flow of the micelles. As shown in Figure 1.4, micelles are also swept toward the negative electrode, with analyte, and past the detector. SDS was the surfactant used in all MECC experiments discussed in this thesis.

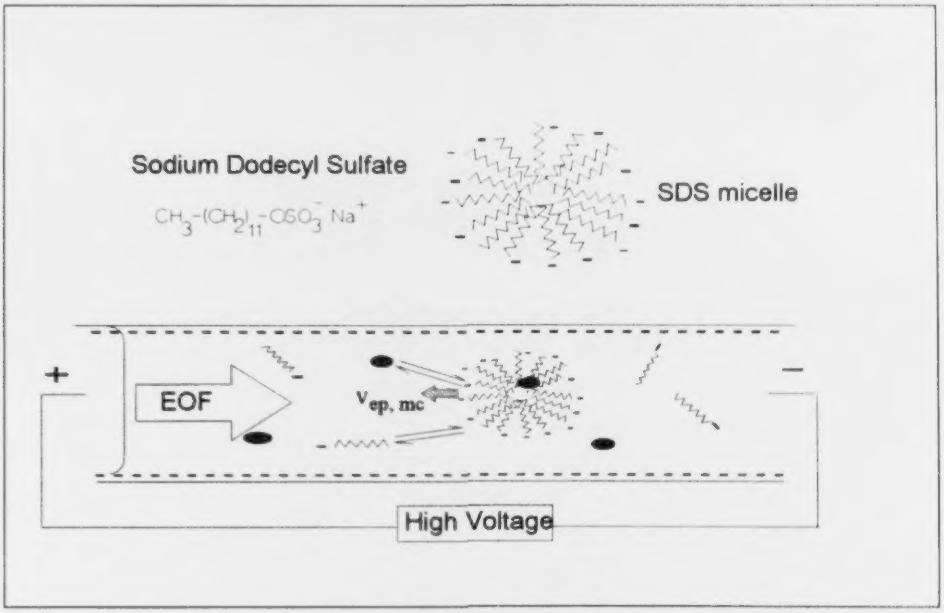


Figure 1.4 Micellar electrokinetic capillary chromatography using SDS.

Several examples of non-ionic, cationic and zwitterionic surfactants exist for MECC separations of various molecules [13-16]. Solute migration in MECC is similar to that of CZE. While there is an additional contribution to bandbroadening, namely resistance to mass transfer in the pseudo-stationary phase, this effect is small and the same equations can be used to estimate N. The resolution in MECC [12] is expressed as:

$$R = \frac{\sqrt{N}}{4}\left(\frac{\alpha-1}{\alpha}\right)\left(\frac{k_j'}{k_j'+1}\right)\left(\frac{1-{}^{t_0}/_{t_{mc}}}{1+\left({}^{t_0}/_{t_{mc}}\right)k_i'}\right) \tag{1.19}$$

where the separation factor α is equal to k'_j/k'_i which are the capacity factors of the i_{th} and j_{th} analytes. Details for optimizing the migration time, resolution, and resolution per unit time in MECC are discussed in depth by several authors [17-19]. Ionic species, as well as neutrals, are routinely separated by MECC. Depending on their charge, some ionic solutes ion-pair with micelles whereas others partition within. The mechanisms for ionic solute interactions with micelles are not as well established and will not be addressed here.

For some separation applications, the electroosmotic flow (EOF) is supressed by coating the inner walls of the capillary to block silanol groups. This process is briefly described in Chapter 5 where it has been applied.

1.3 Thermo-Optical Absorbance Detection

Detection in Capillary Electrophoresis

Detection of weakly absorbing samples has always been difficult and detection in nanolitre volumes presents a challenge. To preserve the high-speed/high-efficiency characteristics of capillary electrophoresis, a suitable method of detection must be employed. Very small amounts of analyte must be detected in very small volumes. While UV/Visible absorbance, fluorescence, mass spectrometry and electrochemical detection methods are all used with CE quite successfully [4], fluorescence continues to be the most sensitive method of detection for CE. Limits of detection for tetramethylrhodamine labeled amino acids are in the 10^{-21} mol range [20]. For molecules that don't naturally fluoresce or are cumbersome to label with a fluorophore, UV/Visible absorbance methods are the most routinely used.

All commercially available CE instruments have conventional on-column UV/Visible absorbance detectors; they are simple and inexpensive. Unfortunately, these detectors are not highly sensitive because absorbance depends on optical path length and in capillaries, the path length is very short. Light passing through an absorbing medium loses intensity at a rate $d\mathrm{I}/dl$ proportional to both the absorbing species' concentration, C, and the instantaneous intensity, I, such that:

$$-\frac{d\mathrm{I}}{dl} = \beta \mathrm{I} \mathrm{C} \tag{1.20}$$

where β is a proportionality constant called the absorption coefficient and l is the path length. Rearranging Equation 1.20 and integrating over path length gives the Beer-Lambert Law expressing light transmission through an absorbing sample:

$$\mathrm{I} = \mathrm{I}_0 e^{-\beta l \mathrm{C}} = \mathrm{I}_0 10^{-\varepsilon l \mathrm{C}} \tag{1.21}$$

where $\varepsilon = \beta/2.303$ and is called the molar absorptivity. The exponent, $\varepsilon l \mathrm{C}$, is often simplified to absorbance, A. The absorbance is therefore related to transmittance, T, by:

$$\mathrm{A} = -\log_{10}\mathrm{T} \quad \text{where} \quad \mathrm{T} = \frac{\mathrm{I}}{\mathrm{I}_0} \tag{1.22}$$

Initially, modified HPLC detectors were used for CE; however the detector volume was usually too large. Essentially, the illumination slit width was much larger than capillary inner diameters. Walbroehl and Jorgenson [21] built a UV absorption detector by simply making a smaller pinhole for the lamp. Good linearity was observed with detection limits of 15 pg isoquinoline and 250 pg lysozyme in 75-μm-ID capillaries. A more complicated variable wavelength detector operating at $\lambda = 193$ nm was used for on-column absorbance measurements of proteins separated in a 75-μm-ID capillary [22]. Minimum detectable concentrations for various proteins were about 5×10^{-4} M.

Bruin *et al.* [23] evaluated several arrangements for UV absorbance detection in CE. A focusing lens, rather than slit or pinhole, in front of the capillary resulted in higher sensitivity. One obvious method of increasing detection sensitivity is to increase optical path length. To maintain the high resolution separation capability of narrow bore capillaries, illumination along the column axis is necessary. Z-shaped and U-shaped detection cells showed about four times better signal-to-noise ratios [23] compared with the focusing lens type detector. In another effort to increase optical path length without increasing cell volume, Wang *et al.* [24] designed a multireflective absorption cell. A 40-fold improvement in sensitivity was achieved compared to a single-pass cell. A concentration detection limit of 6.5×10^{-8} M was obtained for brilliant-green dye.

Yeung [25] described an absorbance detection technique whereby the excitation beam was directed along the axis of the entire capillary. All components contribute to absorption initially, and, as they elute out of the column one by one after separation, transmitted intensity steps up for each component. Each step height is equivalent to the integrated peak area. Regardless of the detector configuration, researchers use both conventional light sources (lamps) and lasers in absorbance-based measurements. Each year new schemes are developed to increase detection sensitivity while preserving high efficiency separation in CE.

Photothermal Methods

The use of lasers in microchemical analysis continues to grow, especially as lasers become more reliable and more affordable. The high radiance, spectral purity, and temporal and spatial coherence of lasers make them an ideal optical source in the analytical chemistry laboratory [26]. Lasers that emit high energy picosecond pulses take advantage of temporal coherence, whereas spatial coherence allows the laser beam to be focused to spot sizes approaching the wavelength of light without loss of power.

For probing properties of analyte in the ultra-small volumes of capillary electrophoresis, lasers are nearly perfect. Laser beams can easily be directed to the target and focused down to a small spot with a simple lens. Conventional optical transmission methods for determining analyte concentration are path length dependent. For very short path length samples, a better technique for measuring analyte quantity via absorbance of light is thermo-optical absorbance, a method of photothermal laser spectroscopy [27].

Photothermal methods involve the measurement of heat produced from radiationless relaxation after absorption of a photon. The emitted heat can be measured directly by measuring the temperature change: $\Delta T \approx 10^{-5}$ to 1 K at ambient temperature. Alternatively, the emitted heat can be measured indirectly by monitoring some property of the sample that changes with temperature like density, conductivity or refractive index. Photoacoustic spectroscopy [28, 29] involves the measurement of the pressure wave caused by an acoustic wave generated from thermally induced density changes in a sample. Photothermal conductivity modulation [30, 31] involves the measurement of resistance changes generated from thermally induced electrical conductivity of a sample. Thermo-optical spectroscopy involves the measurement of deflected, defocused or refracted light after passing through a thermally modulated refractive index gradient in a sample. All photothermal techniques respond to very small relative temperature changes, therefore modulated light sources are used for excitation followed by integrated, averaged or phase-sensitive detection of the resultant signal.

The thermal lens effect described by Leite *et al.* [32] in 1964 is a classical example of a thermo-optical technique where a laser beam becomes defocused after passing through a refractive index gradient. Weakly absorbing samples placed inside a He-Ne laser cavity act as a divergent lens because absorption of the laser beam by the transparent liquid causes localized heating and formation of a refractive index gradient. Absorbances around 2 to 6×10^{-4} were calculated for benzene, CS_2, toluene, nitrobenzene and CCl_4. No tabulated absorbances of these materials were previously recorded [32]. Hu and Whinnery [33] noted that thermal lens techniques could probe a refractive index gradient caused by a temperature change of 10^{-5}°C and absorbances as small as 10^{-5} cm^{-1} were measured for weakly absorbing liquids and solids. They measured absorbances of several organic solvents using an extracavity thermal lens. Long *et al.* [34] also applied the extracavity thermal lens method to measure absorptivities of 10^{-6} $Lmol^{-1}cm^{-1}$. Absorbance detection limits of 7×10^{-8} cm^{-1} for CCl_4 were reported using time resolved thermal lens calorimetry [35].

Similar single beam thermal lens techniques have been used but with measurement of the beam deflection after passing through a refractive index gradient. Bornhop and Dovichi [36, 37] developed a simple nanolitre refractive index (RI) detector for a fluid filled cylindrical cuvette. Beam deflection and refraction from the induced thermal lens combined with complicated diffraction effects from the tube walls provided intensity-sensitive detection of glycerol solution down to $\Delta RI = 6 \times 10^{-7}$. Bruno *et al.* [38] improved the thermal stability of the RI detector by surrounding the sample tube with a sheath of refractive index matching fluid. A position sensitive detector was used to monitor beam deflection caused by refractive index changes.

The thermal lens is desirable for studying essentially transparent samples because increased laser power produces a proportional temperature rise within the sample volume [39]. The concomitant refractive index change enhances the thermal lens signal; the result is a linear relationship between signal and sample absorbance. A drawback for its use in microchemical analysis is that the thermal lens signal is, to a first approximation, proportional to path length. The further the laser beam travels through the sample, the more it interacts with the thermal lens producing a larger signal. For analysis of small volumes of sample or short path length samples, the crossed-beam thermal lens is preferred [40, 41].

Crossed-Beam Thermal Lens

In the crossed-beam technique the thermal lens is produced by one laser—the pump beam—and the thermooptical signal from a defocused beam passing through the lens is produced by a second laser—the probe beam. The two beams are coplanar, directed at right angles to each other and focused on the sample as shown in Figure 1.5. The pump and probe beams interact only at their intersection, comprising a volume of a few picolitres. A cylindrical thermal lens is formed rather than a spherical thermal lens as produced by the single beam experiment. The spatial coherence of the laser beam allows it to be focused to a small spot. This characteristic aids in studying small volume samples because signal is inversely proportional to spot size of the pump laser beam [42]. In Figure 1.5, the cylidrical thermal lens is detected by a photodiode, in the far field, as a uniform change in intensity. A typical pump laser is the argon ion laser (λ=488 nm) and a typical probe laser is a low power helium neon laser (λ=633 nm).

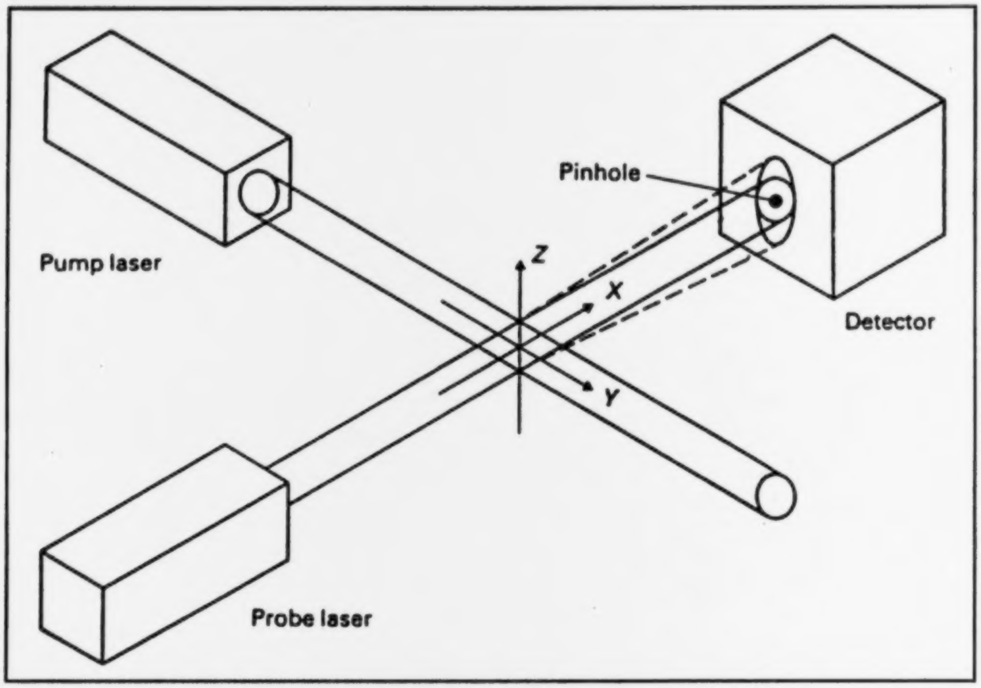


Figure 1.5 The crossed beam thermal lens. The pump beam propagates along the y axis, the probe beam, along the x axis and refraction occurs in the z direction at the intersection volume of the two lasers. The thermal lens signal is measured as an intensity change using a photodiode [39].

The crossed-beam thermal lens technique is also called photothermal refraction, since the probe beam is refracted by the thermal gradient formed during absorption of the pump beam. Since the temperature increase is proportional to sample absorbance, the change in refractive index with temperature is also proportional to sample absorbance. Sample absorbance is proportional to sample concentration, so photothermal refraction is essentially absorbance detection. While the thermal lens signal is proportional to the sample's absorbance of probe laser photons, it is not dependent on the sample path length. The temperature rise in a static, homogeneous sample produced by a pulsed, Gaussian pump beam can be expressed as:

$$\Delta T_{\text{Impulse}}(x,z,t) = \frac{2.303\mathrm{E}\varepsilon\mathrm{C}}{2\pi \mathrm{k} t_C\left(1+2t/t_C\right)} \exp\left\{\frac{-2(x^2+z^2)/\omega^2}{\left(1+2t/t_C\right)}\right\} \qquad (1.23)$$

where t is the time after the pump laser pulse, E is the laser energy, ε is the sample molar absorptivity, C is the sample concentration, k is the thermal conductivity, ω is the pump beam spot size and t_c is the time constant characterizing the temporal response as defined below in Equation 1.24. The numerator in the exponent defines the volume of intersection of the pump and probe laser beams in an *x-y-z* coordinate system [27]. This equation applies for a sample container that has flat walls and does not interfere with the laser beams. Equation 1.23 was constructed using an impulse response function which models pulsed excitation. The time constant expression is:

$$t_c = \frac{\omega^2 C_p \rho}{4\mathrm{k}} = \frac{\omega^2}{4D} \qquad (1.24)$$

where C_p is the heat capacity, ρ is the density and D is the thermal diffusivity.

In many experiments it is more convenient to use a mechanically chopped continuous wave (cw) pump laser rather that an expensive pulsed laser. Convolution of the impulse response function with the appropriate excitation function, e.g. a trapezoidal-based function, is used to obtain a response function for a chopped pump beam. A simple approximation to this waveform is a step excitation function. Equation 1.25 describes the temperature rise in a static, homogeneous sample produced by a cw pump beam turned on at t=0 and having an average excitation power, P:

$$\Delta T_{\text{Step}}(x,z,t) = \frac{2.303\mathrm{P}\varepsilon\mathrm{C}}{4\pi\mathrm{k}}\left\{\mathrm{Exp}_1\left[\frac{2(x^2+z^2)/\omega^2}{1+2t/t_C}\right] - \mathrm{Exp}_1\left[2(x^2+z^2)/\omega^2\right]\right\} \qquad (1.25)$$

where Exp are the exponential decay functions. Equation 1.25 indicates that the temperature rise within a static sample increases linearly with pump laser power per unit area ($P/\pi\omega^2$) and therefore the sensitivity of the thermooptical effect increases. Hence, to improve sensitivity a higher power pump laser could be used or the beam could be focused to a smaller spot.

In separation experiments the sample is flowing rather than static, therefore a model of temperature distribution in flowing samples is necessary. Ignoring thermal gradients generated by flow, the temperature rise in a flowing, homogeneous sample produced by absorbance of a pulsed Gaussian pump beam is:

$$\Delta T_{\text{Flow}}(x,z,t) = \frac{2.303 P \varepsilon C}{2\pi k t_C (1 + 2t/t_c)} \exp\left\{ \frac{-2\left[x^2 + (z + Vt)^2\right]/\omega^2}{(1 + 2t/t_c)} \right\} \tag{1.26}$$

where V is linear flow velocity, assumed to be constant across the beam intersection volume [43]. Equation 1.26 applies to low flow rates directed along the z axis and through a sample cell similar to the static experiment. Equation 1.26 is derived for the impulse excitation response function.

The probe laser, directed at right angles to the pump laser, interrogates the temperature gradient and, hence, refractive index gradient generated by the pump beam. The refractive index change, Δn, is related to temperature rise, ΔT, by:

$$\Delta n = {}^{dn}\!/_{dT}\, \Delta T \tag{1.27}$$

While dn/dT varies from solvent to solvent with temperature, wavelength and pressure, it is considered constant for the small temperature changes produced in the crossed-beam thermal lens and for the narrow bandwidths of the probe laser. If dn/dT varies considerably over the temperature rise then it must be integrated from initial temperature, T_0, to ΔT according to Equation 1.28:

$$\Delta n = \int_{T_0}^{T_0+\Delta T} ({}^{dn}\!/_{dT})\, dT \tag{1.28}$$

As shown in Figure 1.5, signal is detected as a change in the probe beam centre intensity at some distance (far-field approximation) from the laser beams' intersection.

Since the probe beam is aligned on axis at the minimum of the temperature gradient, the probe beam is defocused equally on each side.

A similar experiment to photothermal refraction is photothermal deflection. The probe beam is directed slightly off axis at the maximum of the temperature gradient and thereby deflected away from the heated region. A position sensitive detector is used to measure the change in position of the probe beam [43]. The deflection angle of the probe beam depends on the optical path length through the refractive index gradient according to Equation 1.29:

$$\Phi = \frac{dn}{dT}\int_{path}\frac{\partial \Delta T}{\partial z}dX \tag{1.29}$$

The maximum deflection angle occurs when the probe beam is offset from centre (on axis) by a distance equal to one half the probe beam waist spot-size. It is important to remember that, in flowing samples, the cylindrical thermal lens formed by the pump beam absorbance is distorted by removal of heat.

Capillary Thermo-Optical Absorbance

A combination of the crossed-beam thermal lens technique and the refractive index detector described by Bornhop and Dovichi [37] led to development of an on-column, thermo-optical absorbance detector for capillary chromatography and capillary electrophoresis [44-49]. Figure 1.6 shows the typical arrangement of components in the thermo-optical absorbance detector. A pump laser beam is focused on, and at right angles to, a cylindrical capillary tube, which is shown in cross-section. The pump beam is modulated by a mechanical chopper, providing a reference frequency to the lock-in amplifier. A probe laser beam, coplanar to the pump beam, is focused at right angles to both the capillary tube and pump beam in an x-y-z configuration. Thermo-optical absorption of the sample results in movement of the probe beam, which is detected by the photodiode as a change in beam intensity. The modulated signal intensity, equal to the frequency of the chopper, is measured using the lock-in amplifier. Figure 1.7 shows a more detailed diagram of the pump-probe beam orientation at the capillary.

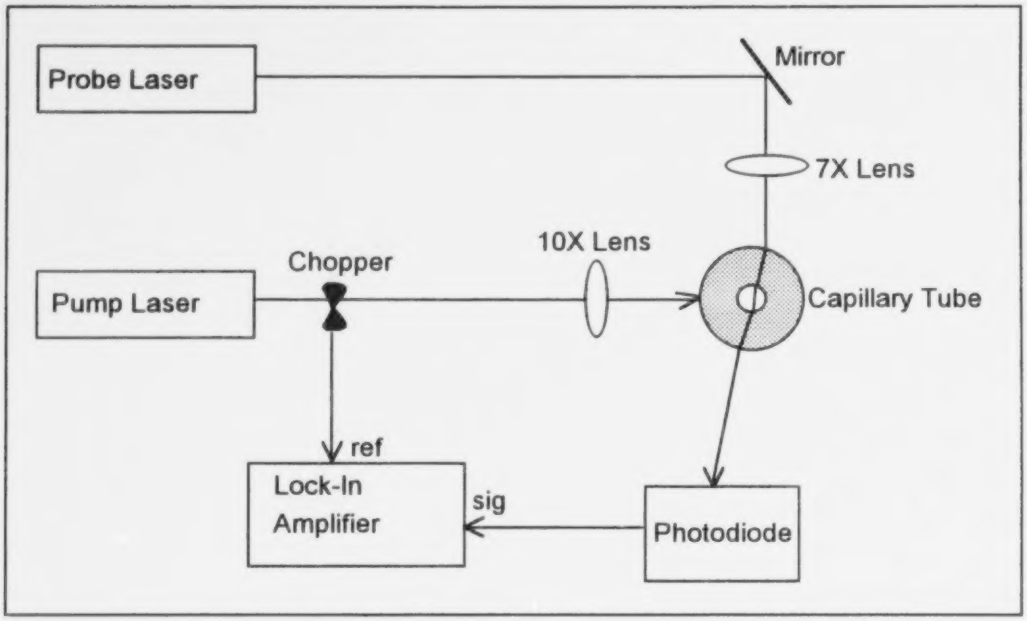


Figure 1.6 Experimental setup for capillary thermo-optical absorbance detector. The mechanical chopper is operated at 50 or 75 Hz. The analog output of the lock-in amplifier is displayed on a strip chart recorder.

Figure 1.7 shows the configuration of the pump and probe beams with respect to the capillary tube. Sample absorption of light from the pump beam followed by radiationless decay of electronically excited states causes localized heating—the photothermal effect. The induced temperature gradient results in a refractive index gradient. The probe laser beam is focused on, and at right angles to, both the capillary tube and pump beam. The probe beam is, though, directed slightly off the axis of the capillary tube centre. Simultaneous diffraction, refraction and deflection of the probe beam are observed because it passes through several mediums of different refractive index: air, then fused silica, then sample solution, and so on. Intensity sensitive rather than position sensitive detection is used. Signal from the photodiode is phase-referenced to the chopping frequency by using a lock-in amplifier. While the modulated component of the probe beam intensity is a measure of sample absorbance, the unmodulated component is proportional to refractive index.

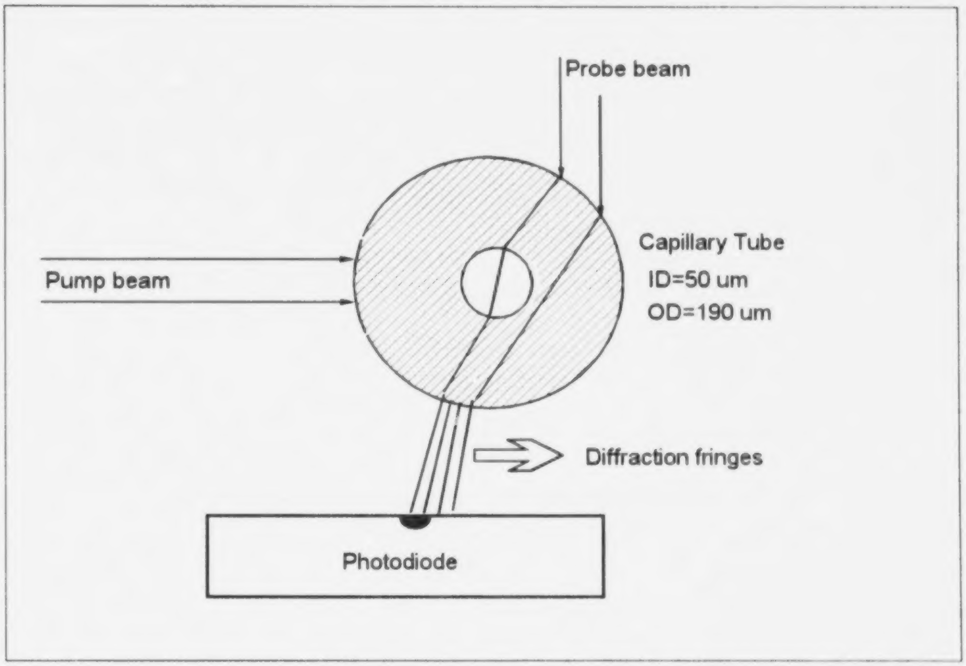


Figure 1.7 Detailed thermo-optical absorbance detection. Movement of diffraction fringes is monitored by the photodiode as a change in probe beam intensity.

The diffraction fringe pattern of the probe beam can be described by conventional light scattering theory for a beam passing through concentric tubes. However, because of the crossed-beam configuration of the experiment photothermal refraction and deflection occur in conjunction with diffraction of the probe beam. In addition, the experiment is designed for chromatographic detection so flowing sample will distort the cylindrical thermal lens by continuous removal of heat in the detection region. Therefore, movement of the fringes is treated phenomenologically. The small area photodiode is positioned beside the first main fringe, which is observed to move when the refractive index of the solution changes. Movement of the fringe results in a change in beam intensity detected by the photodiode. Figure 1.8 is a copy of a photograph of the fringe pattern with the photodiode location superimposed.

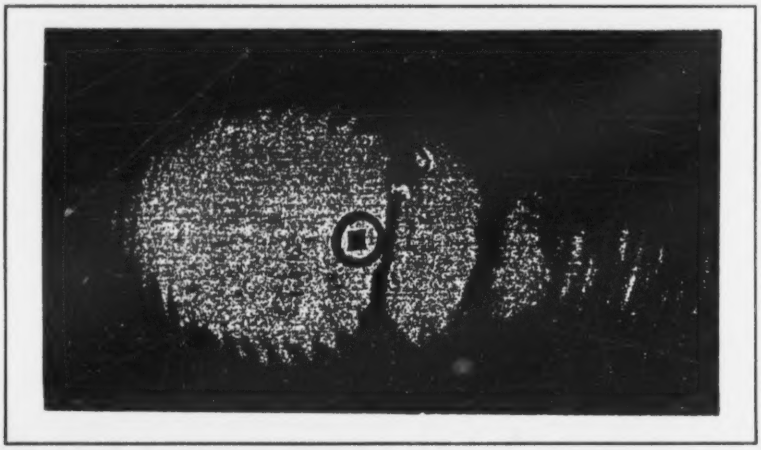


Figure 1.8 Photograph of probe beam profile after diffraction and refraction by sample in the capillary tube. The black circle represents the location of the photodiode with respect to diffraction fringes.

Concentration detection limits (2σ) for static amaranth in methanol solutions were approximately 2×10^{-7} M for each of 50, 100 and 500μm ID capillary tubes [44]. The results indicate that the capillary thermo-optical absorbance signal is independent of optical path length, as in crossed-beam thermal lens techniques. Using the same detection scheme, but for flowing samples separated by capillary electrophoresis, detection limits of about 2×10^{-8} M DABSYL (4,4-dimethylaminoazobenzene-4'-sulfonyl) derivatized amino acids were achieved [45-47]. Within the nanolitre-sized injection volumes, the corresponding mass detection limits are on the order of 40 attomoles. Bruno *et al.* [49] employed the same thermo-optical absorbance detection method to obtain detection limits of 36 pg dansyl (5-dimethylamino naphthalene-1-sulphonyl) derivatized amino acids.

In the work described in this thesis, capillary thermo-optical absorbance detection has been used. In Chapter 2, the use of two different pump laser sources are described. The first is a chopped cw laser and the second is a pulsed laser. Kettler and Sepaniak [50] point out that the crossed-beam photothermal refraction signal will be different for a chopped cw beam versus a pulsed beam. A cw laser has better beam properties and the photothermal signal has a longer rise time than with pulsed lasers. Unfortunately, no inexpensive cw lasers are available in the near UV where many interesting species absorb. Second and fourth harmonics of other laser frequencies will produce cw UV light, but, finicky frequency doublers are required. The main drawback of pulsed lasers is that the peak power density is about 0.2 GW/cm^2, enough incident power to burn a hole through a thin-wall fused-silica flow cell [50]. As a result, moderate power must be used with the sacrifice of some sensitivity; in crossed-beam thermal lens measurements, signal is proportional to pump laser power and inversely proportional to pump beam spot size [27]. In the following chapters, the thermo-optical absorbance detector has been used to identify species separated by capillary electrophoresis.

1.4 Peptide and Protein Sequence Analysis

Terminology

Determination of the complete amino acid sequence of insulin in 1955 by F. Sanger was a landmark in science. Sanger proved that proteins do indeed have unique structures. While biological activity depends on secondary, tertiary and quaternary structure, the primary structure of a protein dictates this active, three dimensional configuration. Knowledge of the amino acid sequence can reveal evolutionary history, elucidate abnormal function and disease and provide access to genes.

A protein is a polymer, comprised of amino acids linked together by peptide bonds—hence the name polypeptide. The primary structure of a polypeptide refers to the sequence in which the amino acids are linked—by convention from the α-amino (N) terminus to the α-carboxyl (C) terminus. Polypeptides of less than fifty amino acids are referred to as, simply, peptides. Large polypeptides that fold and take on tertiary and quaternary structure are referred to as proteins. Most biologically important peptides and proteins contain L-α-amino acids, the structures of which are listed in Appendix A. The twenty amino acids in Appendix A are all the ones that are coded for in eukaryotic DNA. One last note on terminology. An amino acid residue refers to the portion between the peptide bonds—the amino acid minus one water molecule. For example, myoglobin is a protein containing 153 amino acid residues.

Complete analysis of a protein or peptide involves several steps, some of which include isolating and purifying the compound, determining the amino acid composition and sequence, determining its activity via a bioassay, and obtaining its three-dimensional structure via x-ray crystallography and nuclear magnetic resonance (NMR). The work in this thesis focuses on peptide sequence analysis—only one of the important steps in characterizing a protein or peptide.

Primary Structure Determination

The amino acid sequence of a peptide or protein can be determined in two, very different, ways. The first, traditional way to access primary peptide structure is by sequential determination of each amino acid residue of the isolated and purified polypeptide. Sanger [51] first used 2,4-dinitrofluorobenzene (FDNB), which reacts with the amino (N)-terminus of a peptide to form a covalent bond that is stable to subsequent acid hydrolysis. The resultant amino acid derivatized with dinitrobenzene is yellow and therefore can be identified colorimetrically. Other N-terminal reagents include Dansyl chloride and Dabsyl chloride (see Appendix B). Methods for determination of peptide sequences from their carboxyl-terminus are also used but not as routinely and therefore will not be discussed here. The most popular approach to primary structure determination, though, involves the Edman method that was described by Pehr Edman in 1950 [52]. Figure 1.9 shows a very simplified scheme for stepwise sequencing of a purified peptide. In cycle 1, the first (N-terminal) residue is cleaved using Edman degradation chemistry and identified by chromatographic determination. In cycle 2, the truncated peptide from cycle 1 is subjected to Edman degradation to cleave the second residue. The second residue is identified by chromatographic determination. This series of steps is repeated, in a stepwise manner, until the amino acid sequence of the peptide is known. Section 1.4 of Chapter 1 will focus on the Edman method of stepwise degradation for peptide sequencing with brief mention of other technologies where necessary.

The second way to access primary protein structure is from the nucleotide sequence of the gene that codes for the protein of interest. All protein molecules are translated via RNA from the DNA sequence of a particular gene. Advances in technology in the past decade have improved DNA sequencing techniques and now it is a very rapid, accurate method. Knowledge of the genetic code [53] allows one to directly read the corresponding protein primary sequence from the DNA sequence. The main difficulty is in finding the protein-defining sequences in the genome. Even then, the DNA sequence does not always reveal functional relevance without first knowing part of the amino acid sequence of the protein molecule. In many cases, protein sequencing is the only route to the cloning of a gene because it permits design of the oligonucleotide probe required for isolating the gene via recombinant DNA methods [1]. While some may think that knowing a DNA sequence precludes the need for protein sequence analysis, Wittmann-Liebold *et al.* [54] point out that, "since the rapid development of gene technology in the last few years, protein sequence analysis has become even more important".

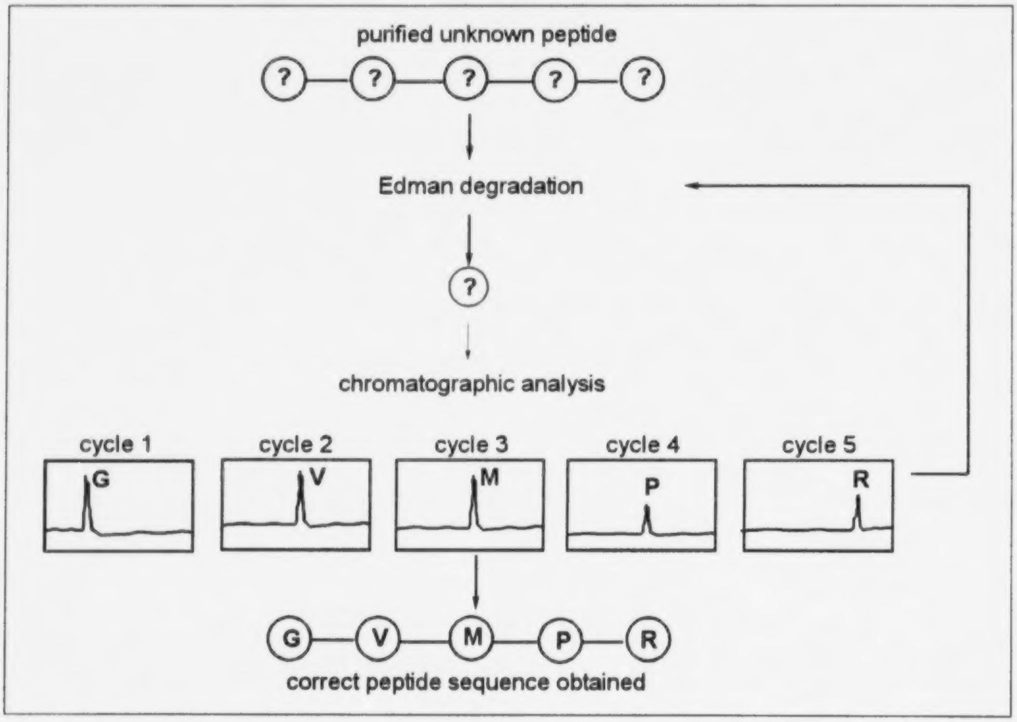


Figure 1.9 Scheme for sequencing a peptide.

An important method of peptide sequencing that deserves mention here, because of its high sensitivity, is direct probe sequence analysis by mass spectrometry (MS) [55, 56]. Several different methods of sample ionization have been used to determine the entire structure of small peptides by MS, but this is a huge field unto itself and will not be discussed further.

Edman Degradation

The Edman method of sequence determination involves reaction of phenylisothiocyanate (PITC) with the N-terminus of a peptide, as shown in Figure 1.10. The first step, coupling, is done under basic conditions at about 50°C. In the second step, cleavage, the resulting phenylthiocarbamyl (PTC) peptide is cyclized under anhydrous acid conditions at about 50°C to produce a thiazolinone (ATZ)-derivatized amino acid (originally the N-terminal amino acid) and a truncated polypeptide. The ATZ amino acid and remaining peptide are separated by organic extraction so the peptide can undergo a second cycle of reaction with PITC, generally refered to as degradation. In the third step, conversion, the ATZ is converted to the more stable phenylthiohydantoin (PTH) form using aqueous acid at about 65°C and identified by chromatography. In 1950 when Edman described this procedure, the PTH derivative was hydrolysed back to the amino acid for identification [52]. The sequencing reaction, which is most often done with phenylisothiocyanate and typically called the Edman degradation, can also be done with other isothiocyanate compounds, which are discussed briefly in this section and in Chapter 2. Details of the reaction chemistry shown in Figure 1.10 are presented in Chapter 3.

The steps involved in the Edman degradation are laborious, requiring several extracting, washing and drying steps and then identification of the PTH amino acids. Croft [59] estimates that in 1960, it took three years and ten grams of material to sequence a peptide having 50 amino acid residues. In 1967 Edman and Begg [60] developed an automated peptide sequencing system that was quickly commercialized and later dubbed a “liquid-phase sequencer”. Compared to manual methods, the automated sequencer was not only faster for sequencing a single peptide but required ten times less material. Reducing the amount of peptide necessary to obtain useful

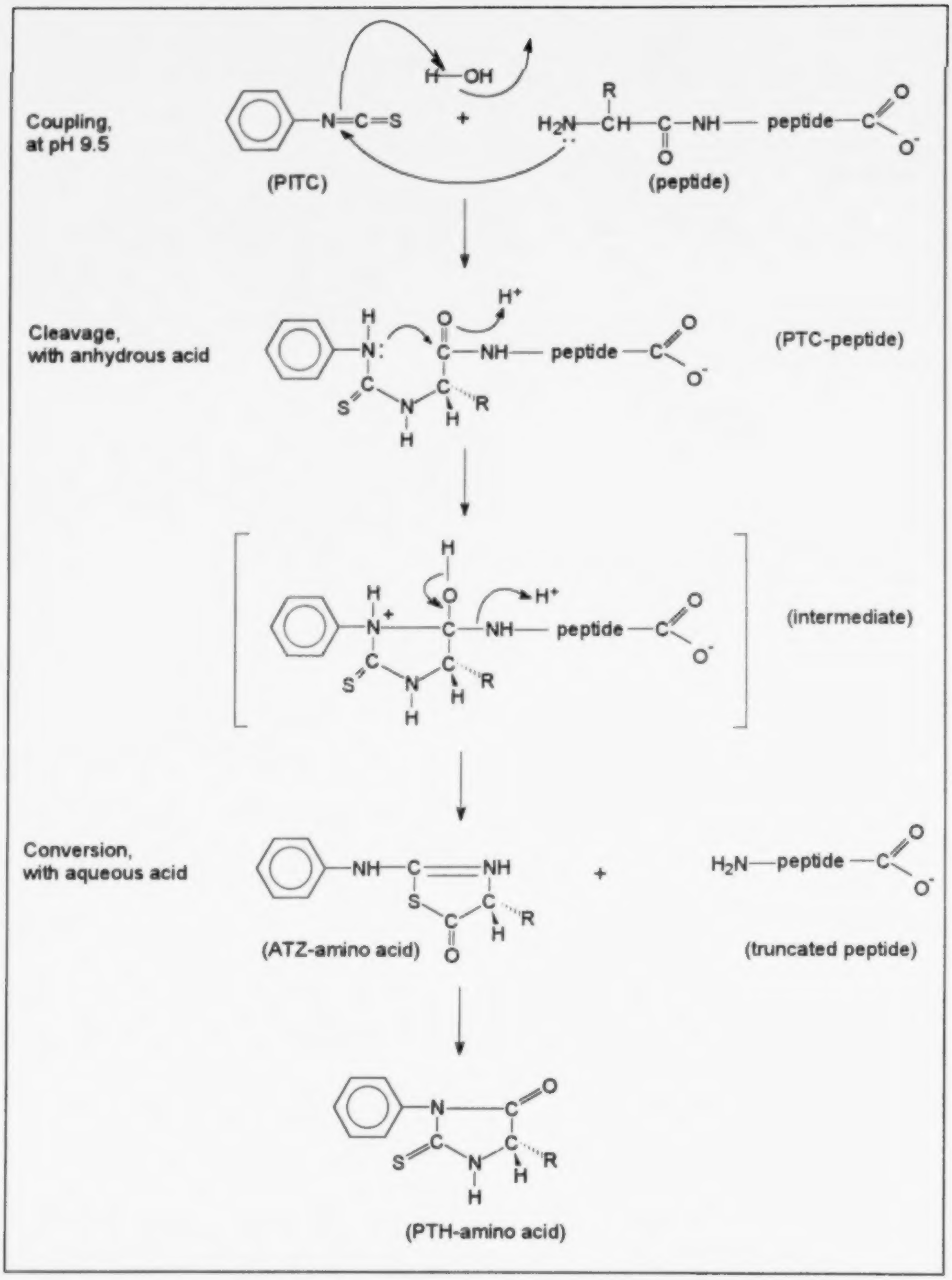


Figure 1.10 The Edman degradation [57, 58].

sequence information is very important since isolation and purification of peptides is time consuming and costly.

Reduction in the amount of peptide needed was, and still is, the impetus for designing better sequencers. In 1971 Laursen [61] introduced a method to sequence a peptide that was covalently attached to a resin: solid-phase sequencing. Reduced sample wash-out from stringent Edman reagents and solvents meant it was now possible to degrade 100 nmol of peptide. One of the major limitations in amount of peptide sequencable was the minimum amount of PTH amino acid that could be visually detected; identification of the PTHs was by thin layer chromatography (TLC) with staining. In an attempt to improve the sensitivity of amino acid residue determination, new degradation reagents were developed.

In 1976 Chang and co-workers [62] introduced dimethylaminoazobenzene isothiocyanate (DABITC) as an N-terminal Edman-type reagent. DABITC amino acid derivatives (DABTHs) are coloured red and are visible in low picomole quantities after TLC separation. Other Edman-type reagents that have been used to sequence peptides include fluorescein isothiocyanate (FITC) [63-66], Boc-aminomethyl-phenylisothiocyanate (BAMPITC) [54], dimethylamino-naphthyl-5-sulfonyl-amino-phenylisothiocyanate (DNSAPITC) [54] and 3-[4'(ethylene-*N,N,N*-thrimethylamino) phenyl]-2-isothiocyanate (PETAPITC) [67]. Determination of DABITC labeled amino acids is covered in detail in Chapter 2. The main reason these highly sensitive derivatizing reagents have not made huge in-roads as routine sequencing agents is because their coupling reaction with the amino-terminus of a peptide is not quantitative. Usually a double coupling step with PITC or methylisothiocyanate (MITC) is required to clean up any unreacted peptide molecules. PITC is still the best Edman reagent with respect to quantitative reaction, even if it is not highly sensitive spectrophotometrically.

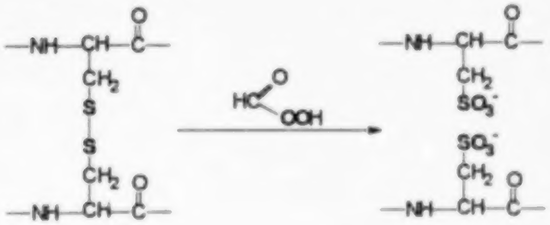
Besides improvements in the Edman chemistry, improvements in instrumentation have been successful in reducing the amount of peptide required for sequencing. Development of the gas-liquid-solid phase peptide/protein sequenator in 1981 [68], followed by improvements in narrow-bore reversed-phase HPLC for PTH residue identification [69], brought routine sequencing down to the picomole level. Unfortunately, many biologically important peptides and proteins are only available in quantities less than one picomole [1, 70] and so researchers are continuing to strive for more sensitive methods of obtaining primary sequence information. To appreciate the motivation behind the research described in this thesis, it is important to understand some of the steps involved in sequencing a protein and, therefore, why the amount of starting material is so critical.

Several steps are involved in obtaining the primary sequence of an unknown peptide or protein. In cases where the polypeptide is short, N-terminal sequencing can be performed after purification. When the protein is large, internal sequencing is performed by fragmenting the large protein into small, more manageable peptides that are then sequenced. Regardless of the approach, some or all of the following seven steps must be performed to assemble the complete sequence of an unknown polypeptide.

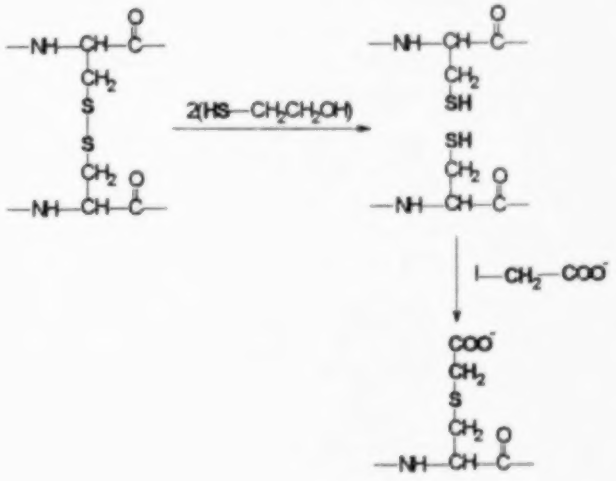
1. The material must be pure. The protein of interest must be isolated from other cellular contents by centrifugation, dialysis, electrophoresis or chromatography or a combination of all of these. An easy way to check purity is by isoelectric focusing (IEF) to see if a single charged species is present or by mass spectrometry (MS) to see if a single molecular mass is present. Also, the biological activity or the property that attracted attention to the protein should remain constant upon further attempts to purify.

2. The material must have unique N- and C-termini. The protein may have more than one chain and/or disulfide bridges. End-group determination (see Appendix

B) or SDS-PAGE (polyacrylamide gel electrophoresis) under oxidizing or reducing conditions will indicate the presence of more than one chain. If so, the polypeptide chains must be separated by chromatographic or electrophoretic means; sequencing will only work if there is a unique N-terminus. Disulfide bridges are broken by oxidation with performic acid to give cysteic acid (Equation 1.30) or by reduction with β-mercaptoethanol and blocking with iodoacetate to give δ-carboxylmethylcysteine (Equation 1.31).



(1.30)



(1.31)

3. The amino acid composition of the polypeptide must be determined. A constant amino acid composition, after further attemps to purify material, will help to confirm its purity. Quantitative amino acid composition will also indicate the size of the polypeptide. Mass spectrometry, another sensitive way to determine molecular mass, can be used to estimate the number of residues. If the chain is much greater than 50 residues, it should be cleaved into smaller fragments. Sequence information beyond about 50 residues is not reliable because of accumulated impurities, incomplete reactions, non-specific cleavages and incompletely extracted ATZ-amino acids.

4. Polypeptide chains may be cleaved into manageable fragments for internal sequence analysis. The purified polypeptide is separated into at least two aliquots and each subjected to a different method of cleavage. Cleavage methods are either enzymatic or chemical and fairly specific as to which peptide bonds are broken. Table 1.1 indicates just a few of the commonly used methods of fragmentation of large polypeptide chains [59]. Native proteins are often not susceptible to enzymatic cleavage and have to be denatured first by heating, acid oxidation or cystine reduction. After fragmentation, the peptides must be separated by chromatographic or electrophoretic methods. Amino acid composition can be performed on a small amount of each peptide to aid subsequent sequence determination. Comparison to the amino acid composition in step 3 provides a material balance.

5. The sequence of separated, purified peptides is determined. Peptides can be sequenced by Edman degradation as described earlier in this section.

6. The peptide sequences are assembled. Overlapping regions of peptides are used to decide the order in which fragments occur from the N- to C-termini in the protein. For example, chymotryptic peptides will show the order in which tryptic peptides must be arranged. Some ambiguities usually exist, but generally only one arrangement of the whole chain is consistent with the sequences of the previously cleaved peptides.

7. Location of disulfide bridges, if any, must be determined. Also, other non-amino acid moieties like covalently bound prosthetic groups or carbohydrates are located. To find disulfide linkages, intact protein must be used. First radiolabeled iodoacetate is used to label free cysteines then the protein is oxidized to break -S-S-bridges. Fragmentation and isolation to create the same peptides as in step 4 will distinguish the bonding from the non-bonding cystines. Next, iodoacetate labeling then fragmentation with the same cleavage reagents without first breaking the -S-S- bridges will give different peptides that, when isolated and sequenced, will show bridge locations. Identification of covalently bound prosthetic groups like heme in cytochrome *c* or carbohydrates in glycoproteins [71] are also necessary but beyond the scope of this paper.

Table 1.1 Methods of fragmentation for large polypeptides.

Reagent	**Cleavage at '—'**	**Exceptions**	**Conditions**
trypsin (enzymatic)	Arg, Lys—X	-when Pro is the next residue.	pH 8-9 37°C
α-chymotrypsin (enzymatic)	Trp, Tyr, Phe, Leu—X, and Met, Asn, Gln, His, Thr—X, but to a lesser extent than top 4.	-when Pro is the next residue.	pH 8-9 37°C
thermolysin (enzymatic)	X—hydrophobic amino acids	X—Leu—Pro	pH 8 40°C
pepsin (enzymatic)	similar to chymotrypsin but less specific, aromatic amino acids—X— —aromatic amino acids Leu—X—Leu		pH 2 22°C
staphylococcal protease (enzymatic)	Glu—X Asp—X some Ser—X	-when Pro is next. Asp-X in carbonate buffer.	pH 4.0 pH 7.8
cyanogen bromide (chemical)	X_1—Met—X_2 to give X_1—homoserine lactone and NH_2— X_2	N-acetyl-Met	low pH 25°C
hydroxylamine (chemical)	Asn—Gly		low pH 45°C
BNPS-skatole (chemical)	Trp—X		low pH 22°C
N-bromosuccinimide (chemical)	Tyr, Trp—X	not as specific as BNPS-skatole	pH 4 22°C

The previous seven steps constitute a basic regime for determining the primary structure of an unknown peptide or protein. Comparison of a short N-terminal sequence with published sequences can often aid positive identification of a protein. There are many cases where N-terminal sequencing is impossible due to a blocked or masked N-terminal amino acid. Initially, lack of sequence information was thought to be because the peptide was cyclic [71]. Later, researchers discovered that a large percentage of proteins simply do not have a free α-amino termial group available for coupling with PITC. The most common block is the N-terminal acetyl group. Another culprit is pyrrolidone carboxylic acid (pyroglutamic acid) formed under slightly acidic conditions from cyclization of an N-terminal glutamine. In these cases, internal sequencing strategies (fragmentation into shorter peptides) are required and/or sequencing from the C-terminus must be performed. Similarly, the location of various reactive groups in a protein may determine its secondary structure such as folding and coiling. These aspects are the essence of 3-dimensional protein structure and the basis of specific reactivity; however they also will not be further addressed in this paper.

The steps above indicate how difficult primary sequence analysis is. While 5 pmol of pure peptide may be sequenced by modern methods, one requires at least 100 times more material to get to the sequencing stage. As mentioned above, many important peptides and proteins are present in biological systems only in trace amounts. To address this problem improvements in sample isolation, sample preparation, sample immobilization, sequencing chemistry, and sequencing instrumentation are continuously being made. The introduction sections of Chapters 3 and 4 go into more detail about these current efforts. In Chapter 3, improvements in identification of PTH amino acids by CE/thermo-optical absorbance detection are discussed. Manual Edman degradation is used to generate the PTH derivatives. In Chapter 4, improvements in sequencing instrumentation are discussed. A miniaturized peptide sequencer is developed to match the volume requirements of CE/thermo-optic determination of PTH amino acids.

1.5 References

1. L.M. Smith, *Anal. Chem.* **60**, 381A-390A (1988).
2. R.A. Wallingford and A.G. Ewing, *Adv. Chromatogr.* **29**, 1-76 (1989).
3. S. Hjerten, *Chromatogr. Rev.* **9**, 122-219 (1967).
4. W.G. Kuhr and C.A. Monnig, *Anal. Chem.* **64**, 389R-407R (1992).
5. K.D. Altria and C.F. Simpson, *Anal. Proc.* **23**, 453-454 (1986).
6. H. Carchon and E. Eggermont, *Amer. Lab.* **January**, 67-72 (1992).
7. J.W. Jorgenson and K.D. Lukacs, *J. Chromatogr.* **218**, 209-216 (1981).
8. C. Schwer and E. Kenndler, *Anal. Chem.* **63**, 1801-1807 (1991).
9. D. McManigill and S.A. Swedberg in *Techniques in Protein Chemistry,* (eds. T.E. Hugli) 468-478 (Academic Press; San Diego, 1989).
10. B.L. Karger and F. Foret in *Capillary Electrophoresis Technology,* (eds. N.A. Guzman) 3-64 (Marcel Dekker, Inc.; New York, 1993).
11. S. Terabe, K. Otsuka, K. Ichikawa, A. Tsuchiya and T. Ando, *Anal. Chem.* **56**, 111-113 (1984).
12. S. Terabe in *Capillary Electrophoresis Technology,* (eds. N.A. Guzman) 65-87 (Marcel Dekker, Inc.; New York, 1993).
13. N. Matsubara and S. Terabe, *Chromatographia* **34**, 493-495 (1992).
14. S.A. Swedberg, *J. Chromatogr.* **503**, 449-452 (1990).
15. Y. Walbroehl and J.W. Jorgenson, *Anal. Chem.* **58**, 479-481 (1986).
16. R.O. Cole and M.J. Sepaniak, *LC-GC* **10**, 380-385 (1992).
17. J.P. Foley, *Anal. Chem.* **62**, 1302-1308 (1990).
18. M.J. Sepaniak and R.O. Cole, *Anal. Chem.* **59**, 472-476 (1987).
19. M.G. Khaledi, S.C. Smith and J.K. Strasters, *Anal. Chem.* **63**, 1820-1830 (1991).
20. J.Y. Zhao, D.Y. Chen and N.J. Dovichi, *J. Chromatogr.* **698**, 117-120 (1992).
21. Y. Walbroehl and J.W. Jorgenson, *J. Chromatogr.* **315**, 135-143 (1984).
22. J.S. Green and J.W. Jorgenson, *J. Liq. Chrom.* **12**, 2527-2561 (1989).

23. G.J.M. Bruin, G. Stegman, A.C. VanAsten, X. Xu, J.C. Kraak and H. Poppe, *J. Chromatogr.* **559**, 163-181 (1991).

24. T. Wang, J.H. Aiken, C.W. Huie and R.A. Hartwick, *Anal. Chem.* **63**, 1372-1376 (1991).

25. E.S. Yeung in *Capillary Electrophoresis Technology,* (eds. N.A. Guzman) 587-603 (Marcel Dekker, Inc.; New York, 1993).

26. D.L. Andrews, *Lasers in Chemistry, 2 ed.*, (Springer-Verlag Berlin; Heidelberg, 1990).

27. N.J. Dovichi, *CRC Critical Reviews in Analytical Chemistry* **17**, 357-423 (1987).

28. E.P.C. Lai, B.L. Chan and M. Hadjmohammadi, *Appl. Spect. Rev.* **21**, 179-210 (1985).

29. E.P.C. Lai, C.S. Silundika, I.W. Wylie and R. Guo, *Spectrochimica Acta Rev.* **13**, 377-398 (1990).

30. R. McLaren and N.J. Dovichi, *Anal. Chem.* **60**, 730-733 (1988).

31. R. McLaren and N.J. Dovichi, *J. Appl. Phys.* **68**, 4882-4884 (1990).

32. R.C.C. Leite, R.S. Moore and J.R. Whinnery, *Appl. Phys. Let.* **5**, 141-143 (1964).

33. C. Hu and J.R. Whinnery, *Applied Optics* **12**, 72-79 (1973).

34. M.E. Long, R.L. Swofford and A.C. Albrecht, *Science* **191**, 183-185 (1976).

35. N.J. Dovichi and J.M. Harris, *Anal. Chem.* **53**, 106-109 (1981).

36. D.J. Bornhop and N.J. Dovichi, *LC-GC* **5**, 427-429 (1987).

37. D.J. Bornhop and N.J. Dovichi, *Anal. Chem.* **58**, 504-505 (1986).

38. A.E. Bruno, B. Krattiger, F. Maystre and H.M. Widmer, *Anal. Chem.* **63**, 2689-2697 (1991).

39. N.J. Dovichi, *Chemistry in Britain* **Sept.**, 895-898 (1988).

40. N.J. Dovichi, T.G. Nolan and W.A. Weimer, *Anal. Chem.* **56**, 1700-1704 (1984).

41. T.G. Nolan, W.A. Weimer and N.J. Dovichi, *Anal. Chem.* **56**, 1704-1707 (1984).

42. T.G. Nolan and N.J. Dovichi, *IEEE Circuits and Devices Magazine* **January**, 54-56 (1986).

43. W.A. Weimer and N.J. Dovichi, *Applied Optics* **24**, 2981-2986 (1985).

44. D.J. Bornhop and N.J. Dovichi, *Anal. Chem.* **59**, 1632-1636 (1987).

45. M. Yu and N.J. Dovichi, *Mikrochim. Acta* **111**, 27-40 (1988).

46. M. Yu and N.J. Dovichi, *Anal. Chem.* **61**, 37-40 (1989).

47. M. Yu and N.J. Dovichi, *Appl. Spect.* **43**, 196-201 (1989).

48. C.W. Earle and N.J. Dovichi, *J. Liq. Chrom.* **12**, 2575-2585 (1989).

49. A.E. Bruno, A. Paulus and D.J. Bornhop, *Appl. Spectrosc.* **45**, 462-467 (1991).

50. C.N. Kettler and M.J. Sepaniak, *Anal. Chem.* **59**, 1733-1736 (1987).

51. F. Sanger, *Biochem. J.* **39**, 507-515 (1945).

52. P. Edman, *Acta Chem. Scand.* **4**, 283-293 (1950).

53. C.K. Mathews and K.E. vanHolde, *Biochemistry*, (The Benjamin/Cummings Pub. Co.; Redwood City, CA, 1990).

54. B. Wittmann-Liebold, J. Shan-Wei and J. Salnikow in *Protein/Peptide Sequence Analysis: Current Methodologies,* (eds. A.S. Bhown) 119-134 (CRC Press; Boca Raton, 1988).

55. H.C. Krutzsch in *Protein/Peptide Sequence Analysis: Current Methodologies,* (eds. A.S. Bhown) 161-180 (CRC Press; Boca Raton, 1988).

56. J.A. Loo, C.G. Edmonds and R.D. Smith, *Science* **248**, 201-204 (1990).

57. P. Edman and A. Henschen in *Protein Sequence Determination,* (eds. S.B. Needleman) 232-279 (Springer-Verlag; New York, 1975).

58. J. McMurry, *Organic Chemistry*, (Brooks/Cole; Monterey, 1984).

59. L.R. Croft, *Introduction to protein sequence analysis*, (John Wiley & Sons; Chichester, 1980).

60. P. Edman and G. Begg, *Eur. J. Biochem.* **1**, 80-91 (1967).

61. R.A. Laursen, *Eur. J. Biochem.* **20**, 89-102 (1971).

62. J.Y. Chang, E.H. Creaser and K.W. Bentley, *Biochem. J.* **153**, 607-611 (1976).

63. H. Kawauchi, K. Muramoto and J. Ramachandran, *Int. J. Peptide Protein Res.* **12**, 318-324 (1978).

64. H. Maeda, N. Ishida, H. Kawauchi and K. Tuzimura, *J. Biochem.* **65**, 777-783 (1969).

65. K. Muramoto, H. Kawauchi and K. Tuzimura, *Agric. Biol. Chem.* **42**, 1559-1563 (1978).

66. K. Muramoto, H. Kamiya and H. Kawauchi, *Anal. Biochem.* **141**, 446-450 (1984).

67. R. Aebersold, E.J. Bures, M. Namchuck, M.H. Goghari, B. Shushan and T.C. Covey, *Protein Science.* **1**, 494-503 (1992).

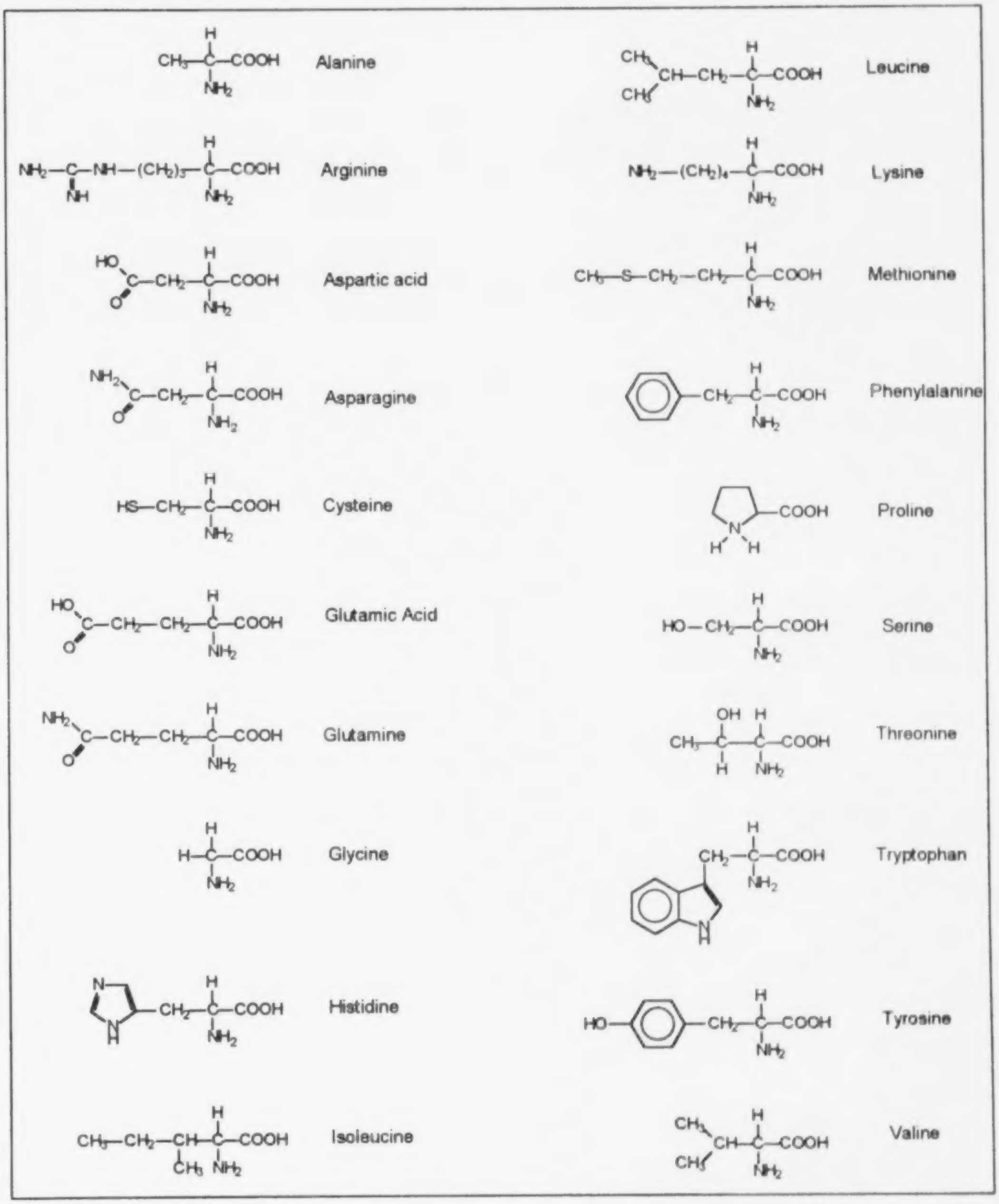
68. R.M. Hewick, M.W. Hunkapiller, L.E. Hood and W.J. Dreyer, *J. Biol. Chem.* **256**, 7990-7997 (1981).

69. R.J. Mattaliano, K.A. Yamada, J.K. Hughes and P.-M. Yuan in *Methods in Protein Sequence Analysis, Proceedings of the 7th International Conference, Berlin, July 3-8, 1988,* (eds. B. Wittmann-Liebold) 112-115 (Springer-Verlag; Berlin, 1989).

70. S. Kent, L. Hood, R. Aebersold, D. Teplow, L. Smith, V. Farnsworth, P. Cartier, W. Hines, P. Hughes and C. Dodd, *BioTechniques* **5**, 314-321 (1987).

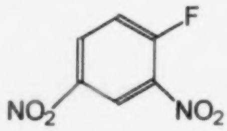
71. W.A. Schroeder, *The Primary Structure of Proteins*, (Harper and Row, Inc.; New York, 1968).

Appendix A: L-α-amino acid structures.

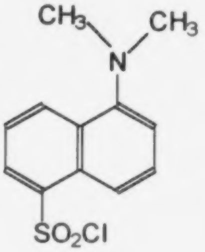


Appendix B: N-terminal derivatizing reagents.

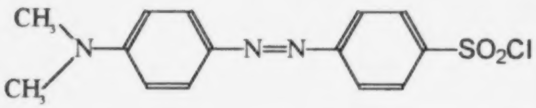
2,4-Dinitrofluorobenzene, Sanger's reagent



Dansyl Chloride



Dabsyl Chloride



CHAPTER 2

CAPILLARY ELECTROPHORESIS
AND
THERMO-OPTICAL ABSORBANCE DETECTION
OF
AMINO ACID DERIVATIVES

2.1 Introduction

The amino acid sequence of a protein is of fundamental value in the study of the structure and function of the molecule. The Edman degradation procedure, described in Chapter 1, is commonly employed to determine the amino acid sequence of proteins [1]. The power of the Edman degradation is demonstrated by its popularity; there are very few routinely used competing technologies for protein sequencing. However, the conventional approach to protein sequencing suffers from several limitations. For example, the macroscopic size of the protein sequencing instruments and the use of conventional liquid chromatography for amino acid detection requires that relatively large amounts of protein or peptide be available for sequencing. Approximately a picomole of purified peptide is required for accurate sequence determination. While mass-spectrometry offers hope for sub-picomole protein sequencing [2], it is important to investigate improvements in current technology.

In this chapter, the use of capillary electrophoresis separation with thermo-optical absorbance detection is investigated as a method to improve the determination of the Edman degradation products: derivatized amino acids. In section 2.2, dimethylaminoazobenzene isothiocyanate (DABITC) derivatives of amino acids are studied. DABITC is an Edman-type degradation reagent offering enhanced detection of amino acid derivatives because of its high molar absorptivity in the visible portion of the spectrum. In section 2.3, phenylisothiocyanate (PITC) derivatives of amino acids are studied. PITC is the traditional Edman degradation reagent used in commercial instruments. Improvements in separation and detection are explored and compared to conventional HPLC in an effort to improve sequencing sensitivity. Section 2.4 of this chapter presents conclusions of the experiments done with DABITC and PITC.

2.2 Dimethylaminoazobenzene Thiohydantoin (DABTH) Amino Acids

To improve detection limits for the determination of cleaved amino acid products, modified Edman degradation schemes have been developed. In each case, the isothiocyanate functional group remains while the phenyl group is replaced with a different chromophore. Perhaps the best developed modified Edman degradation scheme is based on 4-*N,N*-Dimethylaminoazobenzene 4'-Isothiocyanate (DABITC) [3]. This molecule couples with the N-terminus of the polypeptide under basic conditions to produce the thiocarbamyl (DABTC) peptide and cyclizes under acidic conditions to cleave the terminal amino acid as a thiohydantoin (DABTH). Dimethylaminoazobenzene derivatives have a strong absorbance in the blue portion of the spectrum, $\varepsilon \approx 40{,}000$ $\mathrm{Lmol^{-1}cm^{-1}}$, and are usually separated by use of liquid chromatography with absorbance detection at ≈440 nm.

This highly colored reagent produces improvements in the minimum concentration of amino acid detected and hence peptide that can be analyzed. Additional improvements in the mass of protein required for analysis will follow from reduction in the size of sequencing instrumentation. This section of Chapter 2 describes the development of a capillary zone electrophoresis and thermo-optical detection system for separation and identification of minute amounts of DABTH derivatives of amino acids. Capillary electrophoresis offers significant advantages in both sensitivity and speed over liquid chromatography. DABITC, used in a double coupling scheme with phenylisothiocyanate (PITC), could play an important role in the development of an atto-scale protein sequencer.

2.2.1 Experimental

Twenty standard DABTH amino acid derivatives were prepared using a method similar to Chang's [3]. Amino acid (Fluka) stock solutions, 5×10^{-2} M, were prepared in a solution made up of 50 mL acetone, 10 mL water, 5 mL 0.2 M acetic acid, 0.5 mL triethylamine, pH 10.6 and stored at 4°C. Where necessary, the pH was adjusted by addition of 1 M NaOH. A 4×10^{-3} M DABITC (Sigma) stock solution was prepared in acetone and stored at -10°C. To prepare the DABTH amino acid derivatives, 100 μL of each amino acid solution was mixed with 50 μL DABITC stock solution in a 1.5 mL disposable centrifuge vial, placed in a (52 ± 2)°C water bath for one hour, and then dried in a vacuum oven at 30°C. The orange-yellow DABTC amino acid derivatives were acidified with 40 μL water and 80 μL of a 1:2 mixture of 6 M hydrochloric acid/glacial acetic acid. After heating for 50 minutes and drying under vacuum, the red DABTH amino acids were dissolved in 500 μL of a 10 mM phosphate buffer, pH 2.5, to yield a 4×10^{-4} M solution. Figure 2.1 shows the reaction of DABITC with alanine to form DABTH alanine. A sample made up of 30 μL of each of 20 DABTH amino acids, final concentration of 2×10^{-5} M, was prepared in a disposable centrifuge tube.

Capillary electrophoresis of the DABTH amino acids was performed with a 1.1-m long, 50-μm inner diameter fused silica capillary from Polymicro. A mixed 10 mM sodium dodecylsulfate (SDS), 40% acetonitrile, 60% aqueous 10 mM pH 2.5 sodium phosphate buffer was used for the separation. The sample was injected electrokinetically at 20 kV for 5 seconds by replacing the buffer vial with a sample vial, applying voltage, then returning the buffer vial. Thermo-optical absorbance detection was performed 5.5 cm from the ground end of the capillary. The detector has been described in detail in Chapter 1, Figure 1.6, and elsewhere [4, 5]. In this experiment, a 15 mW helium-cadmium laser, $\lambda=442$ nm, provided the pump beam and a 2 mW helium-neon laser, $\lambda=632.8$ nm, provided the probe beam. The pump laser was focused with a 7× objective whereas the probe beam was focused with a 10× objective. A chopping frequency of 84 Hz was employed and the lock-in time constant was 1 second. The analog output of the lock-in amplifier was directed to a strip chart recorder.

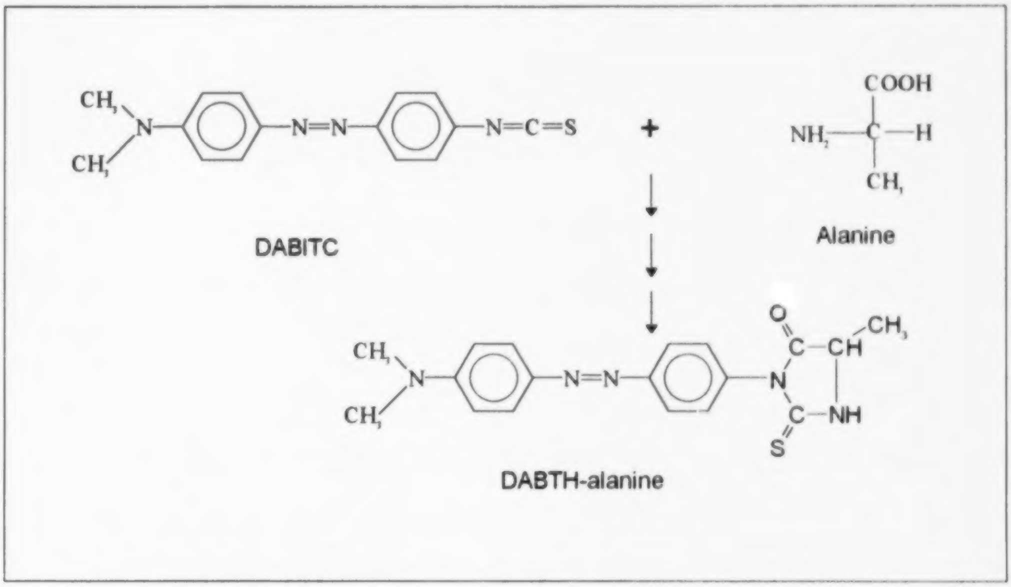


Figure 2.1 Reaction of dimethylaminoazobenzene isothiocyanate (DABITC) with an amino acid.

2.2.2 Results and Discussion

Separation of DABTH amino acid derivatives by electrophoresis is not simple. The structure of a DABTH amino acid is shown in Figure 2.1. The reaction scheme removes both the primary amine and carboxylic acid groups from the amino acid. At neutral pH, the molecules are not charged and therefore have zero electrophoretic mobility. However, under acidic conditions, the secondary amine group does protonate, as it has a pKa of approximately 3. The pKa was determined by monitoring the UV/Visible spectrum of DABTH-tyrosine at various pHs using a Hewlett Packard

8451A diode array spectrophotometer. The protonated and unprotonated forms of DABTH-tyrosine will have different absorbance spectra because these forms have different resonance stabilization. Figure 2.2A shows the absorbance spectrum of the protonated form of DABTH-tyrosine, which exists at low pH. Figure 2.2B shows the spectrum of unprotonated DABTH-tyrosine, which exists at high pH. The higher energy transition (peak) at $\lambda=275$ nm comes from absorption of the aromatic side chain of tyrosine. The lower energy transition (second absorbance peak) at about $\lambda=475$ nm represents absorption of the DABTH portion of the molecule: the portion sensitive to change in pH below 7. This second peak shifts from $\lambda=510$ nm to $\lambda=430$ nm as pH is increased from 1 to 7. A plot of the shift in wavelength for DABTH-tyrosine as a function of pH is shown in Figure 2.2C. In a manner similar to using titration curves, a pKa of 3 for DABTH-tyrosine can be estimated from Figure 2.2C.

The DABTH cations were separated using capillary zone electrophoresis. The resulting electropherogram is shown in Figure 2.3. Careful adjustment of the separation buffer was required to achieve separation of the amino acids. After investigating a number of separation buffers, a 40% acetonitrile, 10 mM SDS/10 mM pH 2.5 sodium phosphate buffer was used to obtain the separation of 15 DABTH amino acids. As in separation of dimethylamino azobenzene sulfonyl chloride (DABSYL) derivatives of amino acids by capillary electrophoresis [4, 5], acetonitrile was added to reduce the zeta potential of the capillary walls. A reduced zeta potential decreases electroosmosis and, therefore, increases the separation time. The electroosmotic flow for this low pH buffer was approximately 0.4 mm/s.

The SDS interacts with the amino acids, perhaps through an ion pairing mechanism, to increase the plate count of the separation. However, separation at low pH is not ideal in fused silica capillaries; relatively low plate counts are produced, 1×10^5 to 4×10^5, possibly due to interaction of the positively charged amino acid derivatives with the negatively charged capillary walls. Detection limits for the separation were on the order of 2×10^{-7} M, similar to reported capillary electrophoresis separation and thermo-optical detection of the DABSYL amino acid derivatives [4, 5].

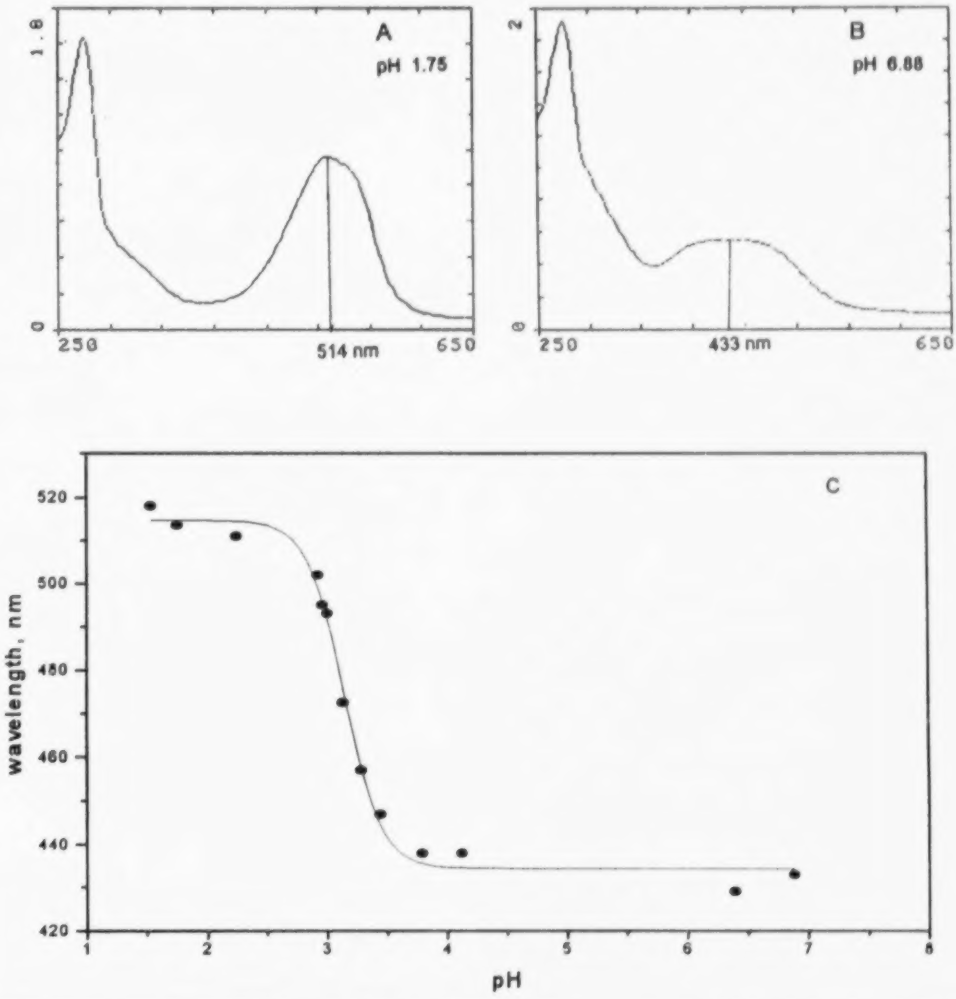


Figure 2.2 Spectrophotometric study to determine the pKa of DABTH-tyrosine: (A) absorbance spectrum of DABTH-tyrosine at pH 1.75, (B) absorbance spectrum of DABTH-tyrosine at pH 6.88, (C) plot of the shift in wavelength with pH for the absorbance of DABTH-tyrosine.

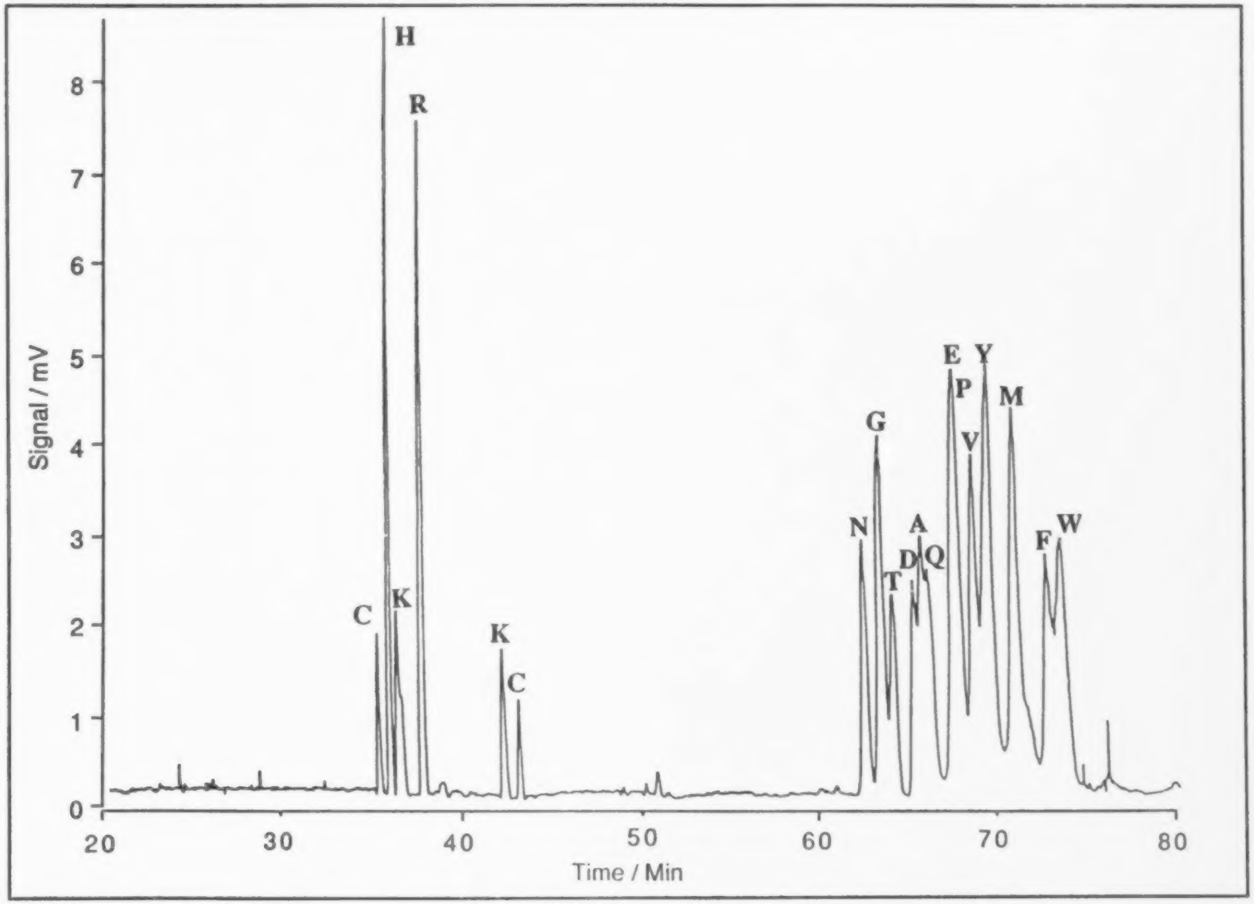


Figure 2.3 Capillary electropherogram of DABTH amino acids. Peaks are indicated by their one letter amino acid codes. Injected concentration is 2×10^{-5} M.

Separation of the DABTH amino acids was not optimized any further for several reasons. First, difficulties with the separation and difficulties with the synthesis of the DABTHs combined to make analysis times very long. A proper vacuum/centrifuge system was not available for evaporating solvents so syntheses took at least two days. Second, stock solutions of DABTH amino acids were not particularly stable, even when stored at -20°C. This instability may be due to slow hydrolysis occuring since the DABTH amino acids were diluted in aqueous solutions. Third, deterioration of the He-Cd probe laser, and eventual complete breakdown, made further experiments impossible. Finally, a low cost UV excimer laser came onto the market that could be used for thermo-optical absorbance detection of commercially available phenylthiohydantoin (PTH) amino acids. PTH amino acids are the products formed in the commercially used sequencing protocol.

2.3 Phenylthiohydantoin (PTH) Amino Acids

As mentioned in section 2.2 of this chapter, alternate isothiocyanates have been proposed for high sensitivity sequencing, including dimethylaminoazobenzene isothiocyanate [3, 6-8] and fluorescein isothiocyanate [9, 10]. While these derivatives may be detected with high sensitivity [11, 12], neither has found acceptance in routine protein sequencing because of inefficient reaction chemistry [13]. To sequence smaller amounts of proteins, improvements are necessary in the determination of PTH amino acids.

Capillary electrophoresis is a useful technique for the rapid and efficient separation of small amounts of analyte [14]. Zone electrophoresis has been used to determine phenylthiocarbamyl (PTC) amino acids; ultraviolet absorbance produces detection limits of ~0.2 picomoles of PTC amino acids [15]. However, zone electrophoresis does not separate the neutral PTH amino acids. Otsuka *et al.* [16] reported the use of micellar capillary electrophoresis (also called micellar electrokinetic chromatography) for the separation of PTH amino acids; distribution of the PTH amino acids between sodium dodecyl sulfate (SDS) micelles and aqueous separation buffer leads to separation. No data was provided on PTH amino acid detection limits, but they are expected to be similar to that for the PTC amino acids. Because of the short optical path length across the capillary, conventional absorbance detection leads to limited sensitivity [17].

Thermo-optical absorbance techniques, discussed in Chapter 1, are well established methods for the determination of small absorbance values [18]. Bornhop and Dovichi [19] developed a thermo-optical absorbance detector for micrometer capillaries used in liquid chromatography. The same detector has been used, with capillary electrophoresis, for detection of molecules that absorb in the visible portion of the spectrum [4, 5, 12, 20, 21]

There are two examples of ultraviolet laser pumped thermo-optical detection for capillary separations. A high energy excimer laser pumped dye laser ($\lambda=400$ nm) was used to determine nitropyrene by capillary liquid chromatography; detection limits were in the mid 10^{-6} M range [22]. Bornhop and coworkers reported the first application of ultraviolet laser-based thermo-optical absorbance detection in capillary electrophoresis [23]. They used a frequency doubled argon ion laser ($\lambda=257$ nm) to determine three DANSYL-amino acids by high speed capillary electrophoresis; detection limits were in the low 10^{-4} M range. Both the excimer laser-pumped dye laser and the frequency doubled argon ion laser are quite expensive and rather temperamental to operate. Neither is likely to see routine use in the analytical laboratory. This section of Chapter

2 describes the use of a low power UV excimer waveguide laser for detection of PTH amino acids after separation by micellar electrokinetic capillary chromatography.

2.3.1 Experimental

A block diagram of the thermo-optical absorbance detector is shown in Figure 2.4. The capillary was illuminated with a 5-mW average power, 10-μJ pulse energy KrF excimer laser (Potomac Photonics Model GX-500) operating at $\lambda=248$ nm, 610-Hz pulse repetition rate, and 50-ns pulse width. The excimer laser beam was focused with a 15-mm focal length quartz biconvex lens at right angles to a 50-μm inner diameter, 190-μm outer diameter fused silica capillary; the polyimide coating of the capillary was burnt from the detection region with a gentle flame. A 3-mW helium neon laser beam (Melles Griot Model 05-LHP-151) was focused at right angles to both the capillary and the excimer laser beam with a 7× microscope objective. The probe beam intensity change was detected 30 cm after the capillary with a 1-mm^2 silicon photodiode. The photodiode output was conditioned with a current-to-voltage converter (1 MΩ feedback resistor in parallel with a 47 pF capacitor) and sent to a two-phase lock-in amplifier (Ithaco Model 3961), phase referenced to the excimer laser pulse repetition rate. An 80386 PC collected data from the lock-in amplifier over the IEEE-488 bus. Sneaker-net was used to transfer the data to a Macintosh IIsi computer. Matlab was used to convolute the data with a Gaussian-shaped filter.

A 39 centimeter long polyimide coated-fused silica capillary was used for this separation; the distance from the injector to the detector was 34 cm. The separation proceeded in a 12.5 mM pH 7.0 sodium tetraborate/sodium phosphate buffer that contained 35 mM sodium dodecyl sulfate (the running buffer). Electrokinetic injection of 5 seconds at 500 V was used. Separation proceeded at 8 kV.

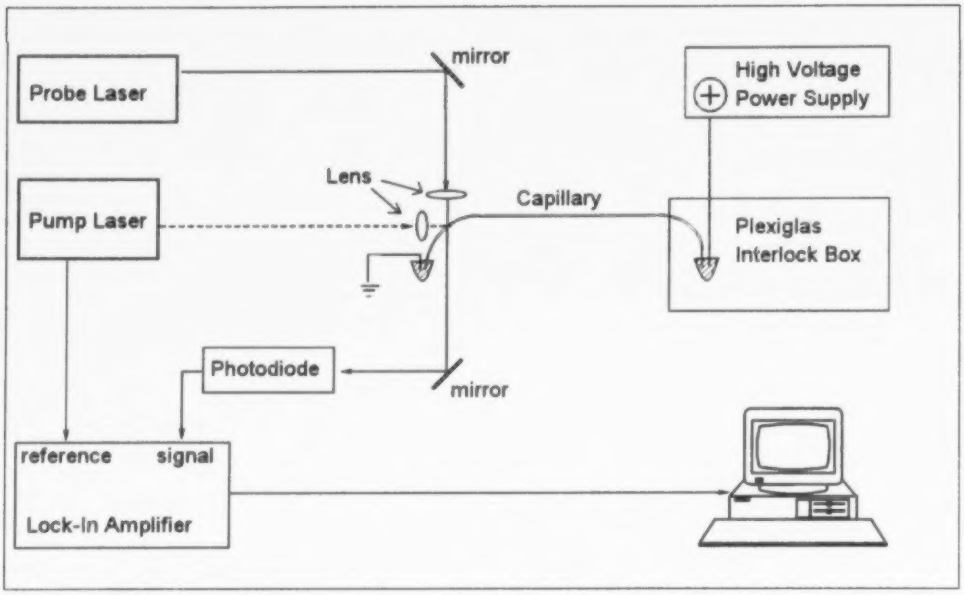


Figure 2.4 Block diagram of experimental setup. A Plexiglas® interlock box is used to protect the user from high voltages during injection and separation of sample. Details are described in the experimental section.

PTH amino acids were purchased from Sigma. Stock solutions of 10^{-2} M were prepared by dissolving each amino acid in a 50% acetonitrile and 50% pH 7.5, 5 mM sodium tetraborate/sodium phosphate buffer. Mixtures of amino acids were prepared by pipetting 10 μL aliquots into 1.5 mL of running buffer.

2.3.2 Results and Discussion

The reaction of phenylisothiocyanate (PITC) with the amino acid alanine to form PTH-alanine is shown in Figure 2.5. Most of the heterocyclic PTH derivatives are neutral at physiological pH and cannot be separated by CZE. It is possible to alter the buffer pH and protonate the PTH derivatives, as described in section 2.2.2 for the DABTH amino acids, to permit CZE separation. However, the addition of surfactant to the buffer to form micelles allows partitioning of the neutral PTH amino acids between the micelles and bulk aqueous phase and permits subsequent separation. This separation method is called micellar electrokinetic capillary chromatograpy (MECC).

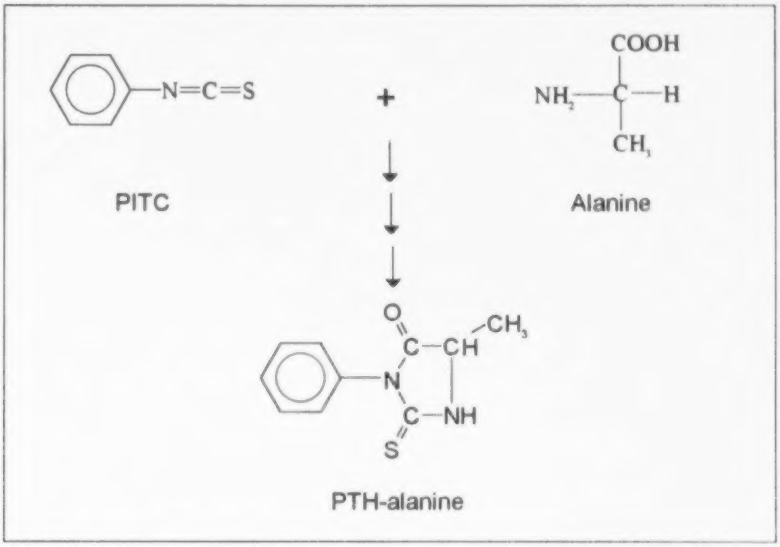


Figure 2.5 Reaction of phenylisothiocyanate (PITC) with an amino acid.

Separation

The electropherogram of 20 PTH amino acids separated by MECC and detected by thermo-optical absorbance is shown in Figure 2.6. A 1.1 mV base line signal is due to absorbance of the excimer laser beam by trace impurities in the separation buffer. The disturbance at 4.5 minutes is due to the elution of trace amounts of acetonitrile, added to the analyte to effect dissolution. Acetonitrile produces a refractive index change that perturbs the optical alignment, producing the base line disturbance. The PTH amino acids are nearly base line resolved, with slight overlap of PTH-alanine and

PTH-glutamic acid and PTH-histidine and PTH-tryptophan. Small signals are typically seen for PTH-serine and PTH-threonine caused by breakdown of these products to their dehydro forms. The separation takes 13.5 minutes at 8 kV.

Separation efficiency in capillary electrophoresis improves with higher applied potentials. Unfortunately, a higher electric field produces significant Joule heating, which generates bubbles in this buffer and destroys the separation. A longer capillary, operated at higher potential (15 kV) but the same electric field (215 V/cm) as the separation represented by Figure 2.6, produces baseline separation of the PTH-amino acids in 40 minutes (Figure 2.7). Peak efficiency is doubled to 416,000 theoretical plates versus 200,000 plates for the separation performed at 8 kV. While peak efficiency is independent of column length as mentioned in Equation 1.8, a longer column permits the use of high separation potential because the current and Joule heating are reduced. Alternatively, Terabe [24] suggested that urea could be used to improve the separation without much penalty in speed. Addition of urea in a few preliminary tests lead to broader peaks and increased background absorbance.

Further optimization of buffer composition and pH for the MECC separation was performed before determining the PTH amino acids generated from Edman degradation. These results are presented in Chapter 3.

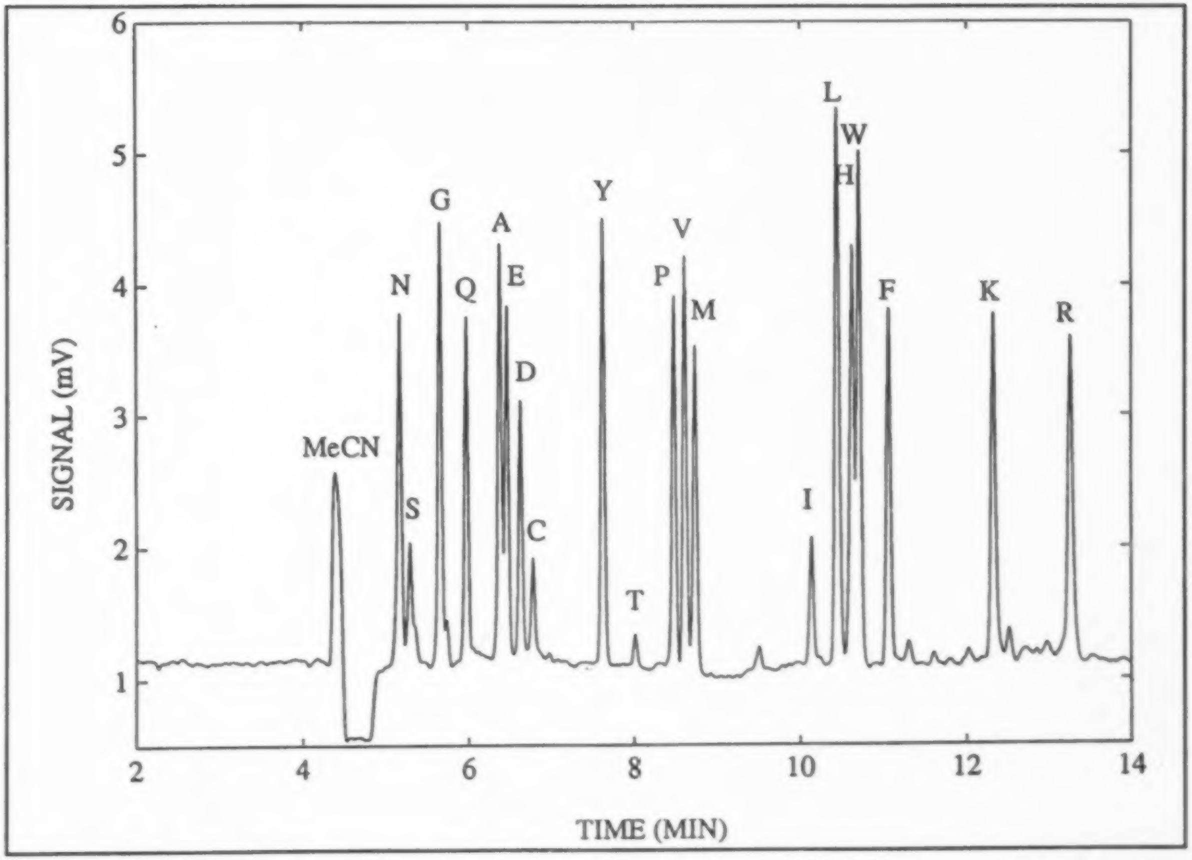


Figure 2.6 High-speed MECC electropherogram of PTH amino acids, 60 fmol each. A 50-µm-ID, 39-cm-long capillary was used for the separation that proceeded at 8 kV in a 12.5 mM phosphate/ borate, 35 mM SDS, pH 7.0 buffer. PTHs are labeled by their one-letter amino acid abbreviations except C, which is cysteic acid.

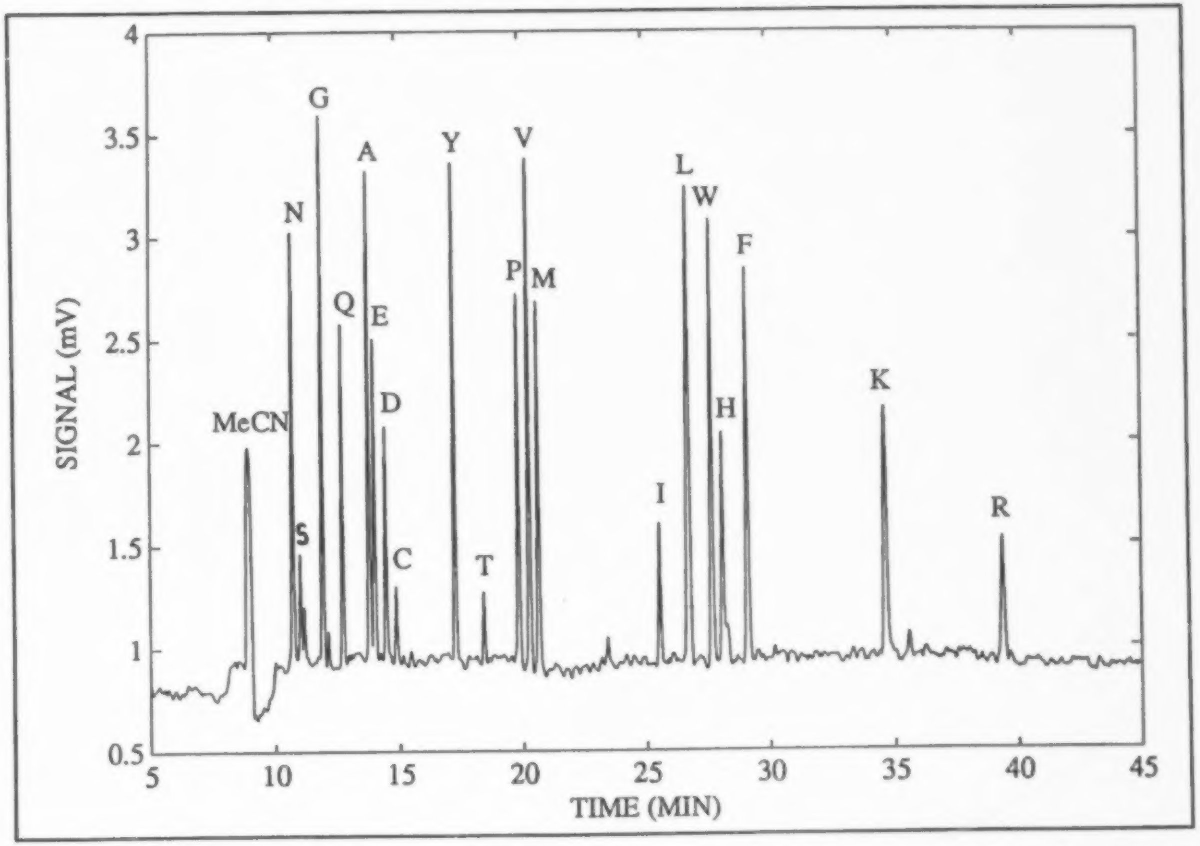


Figure 2.7 High resolution MECC electropherogram of PTH amino acids. A 50-µm-ID, 70-cm-long capillary was used for the separation that proceeded at 15 kV in a 12.5 mM phosphate/borate, 35 mM SDS, pH 7.0 buffer. Injection: 1 kV for 8 s. PTHs are labeled by their one-letter amino acid abbreviations except C, which is cysteic acid. Injected concentration of each PTH amino acid was about 1×10^{-4} M which corresponds to approximately 75 fmol each.

Detection

As noted above, the waveguide excimer laser has a nearly ideal wavelength, consistent pulse repetition rate, good beam spatial coherence, and sufficient pulse energy for thermo-optical detection of PTH amino acids in capillary electrophoresis. At the laser wavelength, 248 nm, the molar absorptivity of PTH-glycine is 6,700 LMol^{-1}cm^{-1}, 60% of the maximum absorptivity at 268 nm (Figure 2.8).

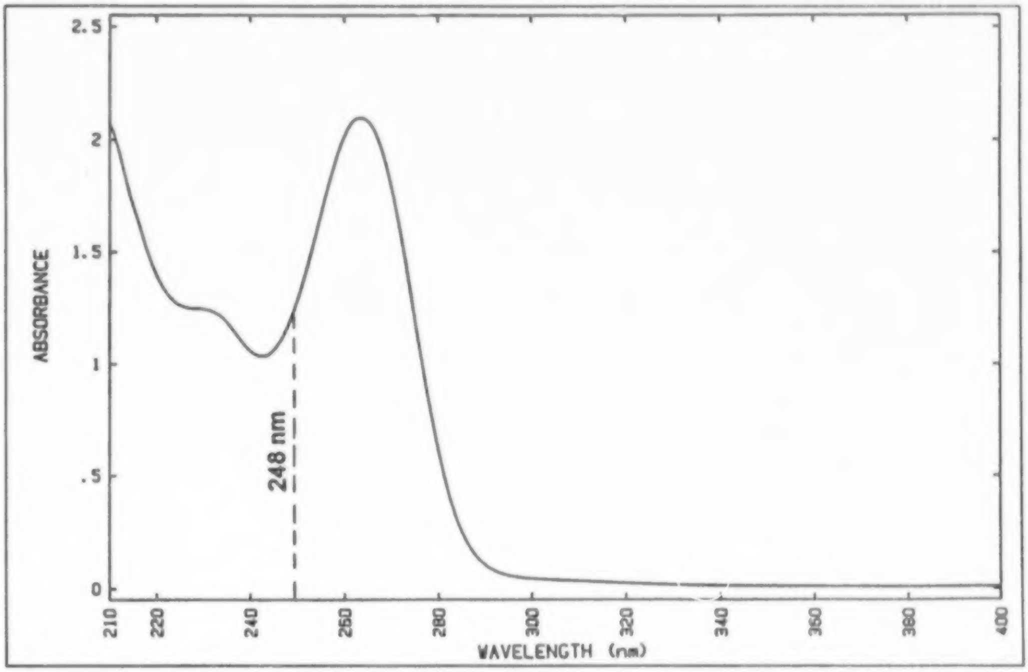


Figure 2.8 Absorbance spectrum of PTH glycine with reference to running buffer, ε_{248}=6,700 Lmol^{-1}cm^{-1} and ε_{268}=11,000 Lmol^{-1}cm^{-1}. Spectrum was obtained on a Hewlett Packard diode array spectrophotometer.

The pulse repetition rate, 610 Hz, was chosen for the best signal-to-noise ratio of the thermo-optical signal. Noise was estimated for several pulse repetition rates by recording signal for two minutes on a strip chart recorder. The 610 Hz rate is well matched to the thermal relaxation time for aqueous solvent in a 50-μm ID capillary. The time-resolved thermo-optical signal was measured at 1 kHz and recorded on a Tektronix digital oscilloscope. The thermo-optical signal is shown in Figure 2.9 and is nearly saw-tooth in shape because most of the signal is generated at the fundamental modulation frequency. As a result, lock-in amplification is quite useful for signal measurement. Because of the waveguide design of the laser, the beam has unusually high spatial coherence for an excimer laser and may be focused easily to a spot size less than the capillary dimensions. Although the thermo-optical signal increases with pump laser power, a modest laser power is appropriate for this experiment as higher energy beams decompose the sample, boil the solution, or destroy the capillary, leading to highly nonlinear response [22]. The peak power for this laser, 200 W, is quite low for an excimer laser because of the modest pulse energy, 10 μJ, and long pulse width, 50 ns.

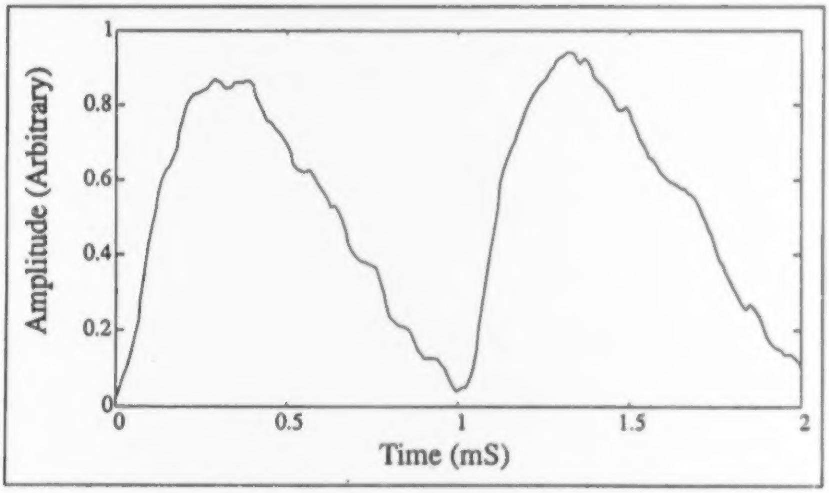


Figure 2.9 Plot of time resolved thermo-optical signal generated by a pulsed excimer waveguide laser operating at 1,000 pulses per second. Two cycles of the waveform are shown.

The time-resolved signal in Figure 2.9 presents a measure of the thermal relaxation rate of solutions in a capillary. The shape of the waveform is dominated by heat flow and not electronic response time. Heat flow from the solution to the capillary is fast and must be accounted for in any model of the temperature rise induced in zone electrophoresis. The origin of the 200-μs rise time is speculative, but may reflect heat flow from the solution to the capillary wall. As shown elsewhere, a parabolic temperature rise is created along the pump beam axis by heat flow to cuvette windows [25]. This temperature rise, and its associated refractive index gradient, acts to defocus the probe beam in the plane containing the pump laser beam.

Detection limit is typically defined as the amount of analyte that generates a signal three times larger than the standard deviation of the base line signal [26]. For 1,000 data points along the baseline, every third point was extracted (to avoid any possible correlation errors) then the standard deviation, σ, was calculated. A calibration curve can then be used to obtain concentration detection limits. The calibration curve for PTH-glycine is shown in Figure 2.10. The slope of the curve is 3.4×10^4 mV/M giving a concentration limit of detection, based on 3σ/slope, of 4×10^{-7} M.

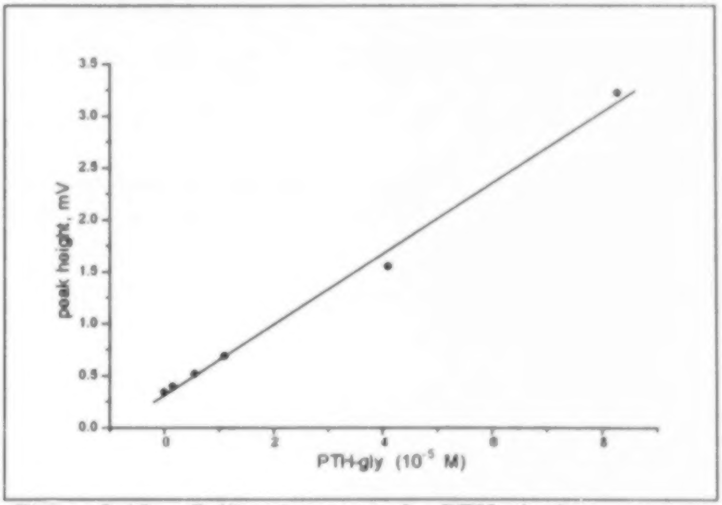


Figure 2.10 Calibration curve for PTH-glycine.

For smoothed data (Figure 2.6), manipulation and detection limit calculation is handled differently. Data in Figure 2.6 were recorded at 0.3-s intervals from a lock-in amplifier that operated with 0.3-s time constant. Peaks are Gaussian in shape with a typical full-width at half-height of 2.5 seconds. Processing of the data was achieved by convoluting the data with a Gaussian filter function with 1-s half-width; noise is reduced and separation efficiency is increased by less than 10%. It is not appropriate to use the standard deviation in the background signal as an estimate of noise for calculating detection limits; the filter function ensures that successive data are correlated. Instead, detection limits can be estimated by Knoll's method for chromatography [27] where the maximum deviation from the mean base line signal, h_n, is measured over a time period given by 50 times the analyte peak width:

$$C_{LOD} = K_{LOD}\, h_n \left(\frac{C_s}{h_s} \right) \tag{2.1}$$

In Equation 2.1, C_s is the amount or concentration of analyte, h_s is the analyte peak height, and K_{LOD} is a multiplier constant derived by Knoll. For a peak-width-multiple of 50, K_{LOD} equals 0.9194 [27].

Calculation of detection limits using Knoll's method is useful when determining several analytes that have different sensitivities. This method precludes the need for twenty different PTH amino acid calibration studies. Detection limits for the 20 PTH amino acids (Table 2.1) range from 0.2 fmol for PTH tryptophan and PTH lysine to 5 fmol for PTH threonine. The volume injected is calculated using the following equation, derived for electrokinetic injection in capillary zone electrophoresis [28]:

$$Vol_i = \frac{(\mu_{ep} + \mu_{eo}) \pi r^2 V_i t_i}{L} \tag{2.2}$$

where V_i is the injection voltage, t_i is the injection time and L is the capillary length to the detector. The moles injected are calculated from the volumes and initial concentrations in the mixture. Poor sensitivity for PTH threonine is due to decomposition of the PTH derivative as stock solutions were stored in aqueous solution leading to long term hydrolysis. These detection limits are two to three orders of magnitude superior to those produced by liquid chromatography and UV absorbance detection [2]. Concentration detection limits, less than 10^{-6} M, are more than two orders of magnitude superior to the other UV thermo-optical absorbance detector in capillary electrophoresis [23]. The use of a high repetition rate excimer laser produces excellent performance.

Comparison of detection sensitivity between the thermo-optical absorbance detector and conventional UV absorbance detectors, which are commonly found in commercial CZE instruments, is more appropriate using absorbance units. One millivolt of thermo-optical signal equals 10^{-4} absorbance units. The base line signal corresponds to an absorbance of 10^{-4} measured across the 50-μm inner diameter capillary. The absorbance per unit length of the electrophoresis buffer is 2×10^{-3} cm^{-1}, a factor of four greater than the (decadic) absorbance per unit length of $(6 \pm 2)\times10^{-4}$ cm^{-1} for water at 254 nm quoted by Boivin *et al.* [29]. Absorbance due to buffer components, and impurities in those components, leads to the observed background signal. The relative noise on the base line, ~0.8%, presumably is proportional to fluctuations in the pump laser pulse energy. The absorbance detection limit (3σ) for the thermo-optical detector is 2.5×10^{-6}. This detection limit is one to two orders of magnitude superior to that produced by commercial transmission detectors in capillary electrophoresis with 50-μm diameter capillaries. Improved detection limits may be produced for the thermo-optical detector by use of higher purity reagents, a more stable pump laser, and solvent with a larger change in refractive index with temperature. While heavy water (D_20) has greater thermo-optical sensitivity than normal water, it has a factor of two greater background absorbance and is not useful as a solvent for high sensitivity ultraviolet absorbance detectors. Mixed acetonitrile-water solvents have a larger dn/dT than water and have proven valuable in capillary zone electrophoresis separation and thermo-optical absorbance detection of DABSYL amino acids [5].

Table 2.1 Detection limits for PTH-amino acids.

PTH Amino Acid	moles injected, fmol	volume injected, nL (± 10%)	detection limit, fmol	detection limit, μM
alanine (A)	66	0.55	0.5	1.0
arginine (R)	22	0.26	0.6	0.9
asparagine (N)	68	0.68	0.7	1.0
aspartic acid (D)	58	0.53	0.8	1.4
cysteic acid (C)	51	0.52	2	3
glutamic acid (E)	65	0.54	0.7	1.2
glutamine (Q)	65	0.59	0.7	1.1
glycine (G)	68	0.62	0.5	0.9
histidine (H)	30	0.33	0.3	0.8
isoleucine (I)	27	0.34	0.7	2
leucine (L)	43	0.33	0.3	0.8
PTH-ε-PTC-lysine (K)	17	0.28	0.2	0.6
methionine (M)	34	0.40	0.4	1.0
phenylalanine (F)	29	0.31	0.3	1.0
proline (P)	41	0.41	0.4	1.0
serine (S)	73	0.66	2	3
threonine (T)	48	0.44	5	12
tyrosine (Y)	42	0.46	0.3	0.7
tryptophan (W)	24	0.33	0.2	0.5
valine (V)	41	0.41	0.4	0.8

An important figure of merit for any analytical technique is the linear dynamic range. The dynamic range of the MECC/thermo-optical detection system was evaluated by injecting a series of PTH-gly samples and is represented by the curve shown in Figure 2.11. The mass detection limit for PTH-gly is 0.2 fmol, based on the calculated concentration detection limit from Figure 2.10 and an injected volume of 0.5 nL. This limit of detection for PTH-gly is slightly better than that presented in Table 2.1, which was calculated from the smoothed data presented in Figure 2.6. Improved optical alignment accounts for the difference in sensitivities.

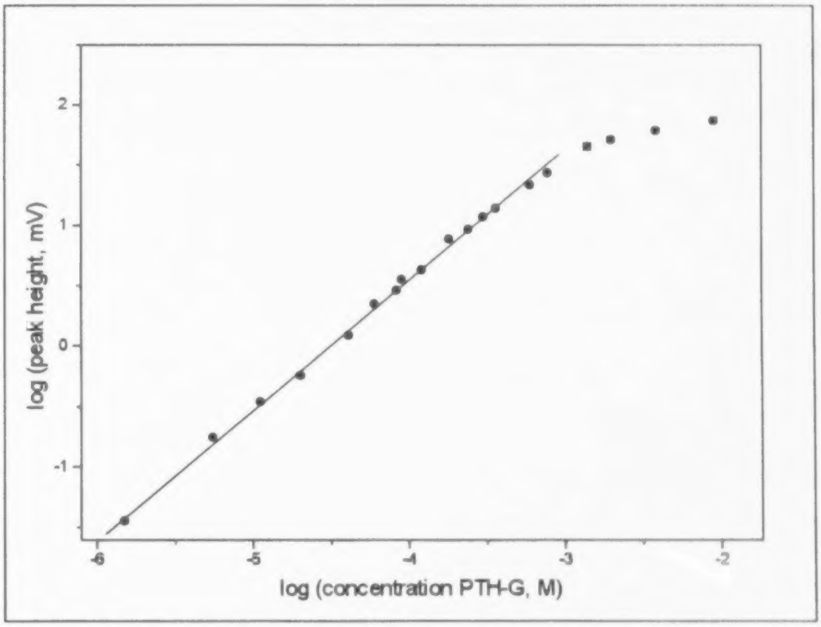


Figure 2.11 Plot of linear dynamic range for PTH glycine.

The slope of the curve in Figure 2.11 equals 1.06, for PTH-gly concentrations ranging from 2×10^{-5} M to 1.4×10^{-3} M. The data plotted in Figure 2.11 include those from the calibration curve (Figure 2.10), therefore the linear dynamic range from detection limit (4×10^{-7} M) to 1.4×10^{-3} M extends a factor of 3,500 in analyte concentration. At higher concentrations, the peak showed pronounced tailing, presumably due to distortions in the local electric field because of the ionic strength of the analyte [30]. However, peak area is conserved and, while not shown here, is linearly related to analyte concentration from the detection limit (0.2 fmol) to at least 5 pmol, a 20,000 range in analyte amount.

Precision was investigated by repeatedly injecting the same sample; the relative standard deviation in peak height was 2.2% for the electrokinetic injection. To obtain this high level of precision, care was taken to ensure that the sample and terminal buffer containers were at the same level. However, signal amplitude is highly dependent on alignment conditions. Replacement of the capillary, and subsequent realignment, inevitably produces a significant change in sensitivity. The capillaries were very stable in this separation and in favorable cases, a capillary would be used for several months before it was replaced. In fact, replacement of the capillary was required only due to breakage, which usually occurred when the fragile detector window was stressed. Migration time precision, of great importance in sequencing, was 1.2%. Both migration time and peak height precision are dominated by reproducibility in adjustment of the power supply voltage for both separation and injection, and in ambient temperature fluctuations.

2.4 Conclusions

CE separation and thermo-optical absorbance detection of DABTH and PTH amino acids were presented this chapter. CE/thermo-optic determination of the Edman degradation products is an important step in the generation of protein sequence information on sub-femtomole samples. In section 2.2, determination of DABTH amino acids was explored as an alternative Edman reagent that has increased sensitivity because of its high molar absorptivity. Also, detection of DABTH derivatives is performed in the visible region of the spectrum (λ=442 nm), unlike the routinely used UV detection of the traditional PTH derivatives. Detection at longer wavelengths should provide better signal-to-noise, eliminating background signal caused by UV absorbing impurities present in the polypeptide degradation chemistry. However, the detection limit of 2×10^{-7} M determined here is only marginally better than previously reported values [4, 5]. Optimization of the capillary electrophoretic separation to give base-line resolution of all twenty DABTHs combined with complete determination in under 25 minutes is necessary in order to pursue experiments of double coupling with PITC or perhaps methylisothiocyanate (MITC) as a clean-up reagent. The method of separation used to determine the DABTH amino acids did not involve micelles but was closer to an ion-pairing method. Perhaps a larger concentration of SDS would improve the resolution of the separation.

Micellar electrokinetic capillary chromatography (MECC) separation with UV thermo-optical absorbance detection of PTH amino acids, presented in section 2.3, is an important step in improving routinely used sequencing methods. Detection is an important issue in capillary electrophoresis [11]. Except for fluorescence [31], most detectors provide limited sensitivity and dynamic range. There is a need for sensitive, universal detection with high dynamic range in capillary electrophoresis. The thermo-optical absorbance detector comes close to satisfying this need; any molecule containing an aromatic group will absorb strongly at this excimer laser wavelength.

MECC separation with thermo-optical absorbance detection in the UV produces outstanding mass detection limits and a wide linear dynamic range for determination of PTH amino acids. Signal strength is, unfortunately, highly dependent on optical alignment and typically must be re-optimized every week. While this alignment sensitivity has little bearing on the qualitative nature of PTH identification for polypeptide sequencing, it hinders precise quantitative measurements. For this reason, a standard mixture of PTH amino acids must be run at the beginning of the day before analysing the sequencing results. The peak heights of the standard PTHs can then be used to quantitate the PTH amino acids recovered from sequencing. Implementation of a servo system to align the system would be an asset. Given these objectives, it is possible to construct a highly miniaturized protein sequencer with a sensitivity of PTH amino acid determination in the sub-femtomole range: at least three orders of magnitude superior to the present state-of-the-art (1 pmol) in amino acid sequencing.

To maintain separation efficiency, the analyte must be injected in nanolitre volumes. As a result, interface with conventional sequencers is problematic. Sample might be concentrated before injection by use of solid-phase extraction, isotachophoresis or stacking [32-35]. Instead, a highly miniaturized sequencer may be developed that is matched to the volume of the capillary electrophoresis separation method. Of course, there are significant experimental difficulties in the design of a miniaturized sequencer. However, because the same reagent concentration is used for the conventional and miniaturized sequencers, the reaction rate and efficiency for the miniaturized sequencer should be identical to, or better than, that of the conventional scale sequencer. Improvements in mass detection limit arise only because the system volume is decreased. Combination of the miniaturized sequencer, as described further in Chapter 4, with capillary electrophoresis separation will allow routine sequencing of femtomole amounts of proteins.

2.5 References

1. P. Edman, *Acta Chem. Scand.* **4**, 283-293 (1950).
2. J.A. Loo, C.G. Edmonds and R.D. Smith, *Science* **248**, 201-204 (1990).
3. J.Y. Chang, *Anal. Biochem.* **102**, 384-392 (1980).
4. M. Yu and N.J. Dovichi, *Mikrochim. Acta* **111**, 27-40 (1988).
5. M. Yu and N.J. Dovichi, *Anal. Chem.* **61**, 37-40 (1989).
6. J.Y. Chang, D. Brauer and B. Wittman-Liebold, *FEBS Letters* **93**, 205-214 (1978).
7. J.Y. Chang, *Methods in Enzymology* **91**, 455-466 (1983).
8. J.-Y. Chang, R. Aebersold, T. Grutter, G. Rosenfelder and D.G. Braun in *Protides of the biological fluids,* (eds. H. Peeters) 955-960 1984).
9. K. Muramoto, H. Kamiya and H. Kawauchi, *Anal. Biochem.* **141**, 446-450 (1984).
10. K. Muramoto, H. Kawauchi and K. Tuzimura, *Agric. Biol. Chem.* **42**, 1559-1563 (1978).
11. S. Wu and N.J. Dovichi, *Talanta* **39**, 173-178 (1992).
12. K.C. Waldron, S. Wu, C.W. Earle, H.R. Harke and N.J. Dovichi, *Electrophoresis* **11**, 777-780 (1990).
13. S. Kent, L. Hood, R. Aebersold, D. Teplow, L. Smith, V. Farnsworth, P. Cartier, W. Hines, P. Hughes and C. Dodd, *BioTechniques* **5**, 314-321 (1987).
14. J.W. Jorgenson and K.D. Lucaks, *Science* **222**, 266-272 (1983).
15. V. Rohlicek and Z. Deyl, *J. Chromatogr.* **494**, 87-89 (1989).
16. K. Otsuka, S. Terabe and T. Ando, *J. Chromatogr.* **332**, 219-226 (1985).
17. J.S. Green and J.W. Jorgenson, *J. Liq. Chrom.* **12**, 2527-2561 (1989).
18. N.J. Dovichi, *CRC Critical Reviews in Analytical Chemistry* **17**, 357-423 (1987).
19. D.J. Bornhop and N.J. Dovichi, *Anal. Chem.* **59**, 1632-1636 (1987).
20. C.W. Earle and N.J. Dovichi, *J. Liq. Chrom.* **12**, 2575-2585 (1989).
21. M. Yu and N.J. Dovichi, *Appl. Spect.* **43**, 196-201 (1989).
22. C.N. Kettler and M.J. Sepaniak, *Anal. Chem.* **59**, 1733-1736 (1987).
23. A.E. Bruno, A. Paulus and D.J. Bornhop, *Appl. Spectrosc.* **45**, 462-467 (1991).
24. S. Terabe, Y. Ishihama, H. Nishi, T. Fukuyama and K. Otsuka, *J. Chromatogr.* **545**, 359-368 (1991).

25. S. Wu and N.J. Dovichi, *J. Appl. Phys.* **67**, 1170-1182 (1990).

26. J.C. Miller and J.N. Miller, *Statistics for Analytical Chemistry, 2nd ed.*, (Ellis Horwood; Chichester, 1988).

27. J.E. Knoll, *J. Chrom. Sci.* **23**, 422-425 (1985).

28. R.A. Wallingford and A.G. Ewing, *Adv. Chromatogr.* **29**, 1-76 (1989).

29. L.P. Boivin, W.F. Davidson, R.S. Storey, D. Sinclair and E.D. Earle, *Applied Optics* **25**, 877-882 (1986).

30. F.E.P. Mikkers, F.M. Everaerts and T.P.E.M. Verheggen, *J. Chromatorg.* **169**, 11-20 (1979).

31. Y.F. Cheng and N.J. Dovichi, *Science* **242**, 562-564 (1988).

32. R.-L. Chien and D.S. Burgi, *J. Chromatogr.* **559**, 141-152 (1991).

33. R.-L. Chien and D.S. Burgi, *J. Chromaatogr.* **559**, 153-161 (1991).

34. D.S. Burgi and R.-L. Chien, *J. Microcol. Sep.* **3**, 199-202 (1991).

35. R.-L. Chien and D.S. Burgi, *Anal. Chem.* **64**, 489A-496A (1992).

CHAPTER 3

CAPILLARY ELECTROPHORESIS
AND
THERMO-OPTICAL ABSORBANCE DETECTION
OF
PRODUCTS FROM MANUAL PEPTIDE SEQUENCING

3.1 Introduction

The sensitivity of peptide/protein sequencing by Edman degradation improved dramatically with the introduction of the gas-liquid-phase sequencer by Hewick *et al.* in 1981 [1]. Sequence information on as little as 5 pmol of protein, a 10,000-fold reduction from ten years earlier, was a remarkable breakthrough. Since the early eighties, several improvements have been made in sample preparation, sequencer instrumentation, and chemistry in an effort to achieve better sensitivity, speed, and ease of operation.

Peptide and/or protein purification is an integral part of microsequencing. Two dimensional gel electrophoresis is a routine analytical tool for purifying proteins where isoelectric focusing in the first dimension is followed by SDS-PAGE (polyacrylamide gel electrophoresis) in the second dimenson. In situ enzymatic digestion of purified protein generates peptides that can be separated and sequenced. Separation of peptide fragments is done either by HPLC or by two dimensional mapping: electrophoresis on a cellulose TLC (thin layer chromatography) plate followed by TLC [2]. The work-up required to prepare a protein for internal sequence analysis is laborious, but improvements in PAGE have resulted in successful sequencing of tryptic peptides obtained from 25 pmol of protein [3]. Advances in microbore liquid chromatography (LC) for fractionation of sub-nanomole amounts of enzymatic digests and peptide purification have also helped to prepare small amounts of sample for sequencing [4].

Electroblotting of proteins from gel electrophoresis onto solid supports that are compatible with gas-liquid-phase sequencing was another important breakthrough in sample preparation for sequencing [5-7]. Utilization of polyvinylidene difluoride (PVDF) membranes [8] and modified PVDF membranes [9-12] for electroblotting has helped reduce the amount of protein needed for sequencing by improving the quantitative transfer of proteins from PAGE.

With respect to instrumentation, advances in column technology have increased the resolving power and reproducibility of RP-HPLC for determination of PTH amino acids [13]. Several improvements in sequencer design, generally towards miniaturization, have contributed to reduced amounts of protein and peptide needed to obtain useful sequence information [14-16]. A relatively new technique for sequencing, which deserves mention, is tandem mass spectrometry (MS) as a direct probe analysis of polypeptides. Sensitivities in MS are about 0.5 pmol [17]; however, applications to peptide sequencing are rather complex. Mass spectrometry is more often used for fast, sensitive molecular mass determination of peptides and proteins than for sequencing.

Improvements in chemistry include efforts to obtain N-terminal sequence information from N-terminal blocked proteins [18, 19], covalent attachment of peptides and proteins to solid supports [20-22], and strategies to identify problem residues such as cysteine [23-26] or identify phosphorylation sites [27].

Despite the developments mentioned above, sequence analysis on polypeptide samples of less than 5 pmol requires heroic efforts. Routine analysis at sub-picomole levels is rare. The main limit is detectability of PTH amino acids—the Edman degradation products. HPLC analysis requires about 1 pmol of PTH amino acid for adequate signal-to-noise and this small amount of product must be dissolved in enough solvent to transfer it to the HPLC. Typical injection volumes for PTH amino acid determination are 50 μL to 100 μL, close to the the upper limit of injection volume to still maintain resonable resolution of closely eluting peaks.

Several methods have been developed to improve sensitivity in detection of PTH amino acids. Much of the work has involved modified Edman-type reagents such as dimethylaminoazobenzene isothiocyanate (DABITC) introduced by Chang *et al.* in 1976 [28], fluorescein isothiocyanate (FITC) [29-32], radiolabeled ^{14}C-PITC or ^{35}S-PITC [33], dimethylaminonaphthyl-5-sulfonylamino phenylisothiocyanate (DNSAPITC) [34] and ethylenetrimethylamino phenylisothiocyanate (PETAPITC) [35]. PETAPITC is an isothiocyanate reagent with enhanced detectability by mass spectrometry. These reagents provide excellent sensitivity for the thiohydantoin amino acid derivatives: 300 amol for P(ETAP)THs [35], 100 amol for DABTHs [36], and 2 zmole for FTHs [37].

The major drawback of homologous Edman-type reagents is their poor coupling chemistry with the N-terminus of a peptide or protein. In most cases, double coupling must be performed due to non-quantitative reaction; each coupling cycle with, for example, DABITC must be followed by a coupling cycle with PITC [34]. This doubles the effort and time of the microsequencing method and effectively halves the number of useful cycles that can be performed due to accummulation of background products in the sequencing reaction. As a result, these reagents have not become popular for routine work. Chemists still prefer to use the tried-and-true phenylisothiocyanate chemistry for sequence analysis. Since PITC degradation works the best for polypeptide sequencing, the efforts of this research project have been directed at improving the sensitivity for identification of PTH amino acids. To this end, I have focused on capillary electrophoretic methods of separation with laser-based detection.

The emergence of capillary electrophoresis (CE) as a routine tool for separating biomolecules has broadened the scope of bioanalytical chemistry [38]. CE offers fast, efficient separation of both charged and neutral species and is rapidly gaining ground in the fields of peptide and protein identification [39-42]. CE is also an excellent method for separating derivatized amino acids [36, 37, 43, 44]. A CE system with thermo-optical absorbance detection was described in Chapter 2 which provides sub-femtomole detection limits for PTH amino acids.

For an analytical system to be useful, it must work with real sequencing samples, not just a purified standard like the the work presented in Chapter 2. This chapter describes the use of CE with thermo-optical detection to identify the products from manual degradation of two peptides: insulin chain B and SP-5 (splenopentin).

3.2 Experimental

Edman Degradation

Sequencing grade ethyl acetate, n-heptane, 12.5% trimethylamine (TMA) in water, Biobrene (Polybrene) and PTH amino acid standard were purchased from Applied Biosystems (ABI, Foster City). Phenylisothiocyanate (PITC) and anhydrous trifluoroacetic acid (TFA), both sequencing grade, were purchased from Sigma as were insulin chain B oxidized and splenopentin (SP-5: Arg-Lys-Glu-Val-Tyr). N-propanol, sodium dodecyl sulfate (SDS), sodium dihydrogen orthophosphate and HPLC-grade acetonitrile were purchased from BDH. Sodium tetraborate and 1-chlorobutane were bought from Fisher Scientific. Pre-purified argon (Union Carbide Canada) was passed through an OT-3-2 oxygen trap purchased from Chromatographic Specialties Inc. (Brockville). Distilled water was further purified through a NANOpure ultrapure water system for buffer preparation.

The chemical protocol used for manual degradation was similar to that used in the Applied Biosystems (ABI) gas-liquid-phase autosequencer, to simulate currently used polypeptide sequencing methods. The reaction mechanism was shown in Chapter 1, Figure 1.10. Reactions were carried out in 600-μL-size polypropylene microcentrifuge tubes. To 150 μL coupling buffer was added 10 to 100 nmol peptide and 15 μL PITC; coupling buffer was 3:2 (v/v) n-propanol:12.5% TMA, adjusted to pH 9.5 with TFA. The solution was flushed with argon, capped, and incubated at 50°C for 30 minutes. For the short peptide (SP-5), 15 μL Biobrene (0.1 mg/mL) was added to the coupling buffer to prevent peptide loss during washing steps.

The resultant PTC-peptide solution was washed by vortexing with 400 μL heptane and the heptane layer discarded after centrifugation. The solution was washed again with 400 μL benzene in the same manner. The aqueous layer was dried on a Savant Speed-Vac for about 1.5 hours. A further wash with sonication using 400 μL ethyl acetate was followed by drying to prepare the PTC-peptide for acid cleavage.

Cleavage involved addition of 50 μL TFA, sonication to dissolve the PTC-peptide, flushing with argon, then incubation at 50°C for 15 minutes. TFA was evaporated in a stream of argon. Extraction of the resulting anilinothiazolinone (ATZ) amino acid was done with 500 μL benzene followed by sonication then centrifugation to carefully separate the organic layer containing ATZ amino acid from the precipitate of truncated peptide. Both phases were thoroughly dried and the peptide subjected to subsequent cycles of coupling and cleavage.

The ATZ amino acid was converted to the more stable PTH-form by addition of 50 μL 25% TFA in water followed by sonication and flushing with argon, then incubation at 50°C for 30 minutes. Residual TFA and water were evaporated and the degradation product (PTH amino acid residue) was stored at -20°C until identification.

Micellar Electrokinetic Capillary Chromatography (MECC)

Determination of PTH amino acids was performed using the CE/thermo-optical absorbance instrument described in Chapter 2, with a few changes: (1) the pump laser was operated at 625 Hz with a neutral density filter (O.D.=0.3) to reduce the beam intensity slightly, (2) the probe beam intensity was detected by a variable gain/variable bandwidth Model 2001 Front-end Optical Receiver (New Focus, Inc), (3) data were collected at 3 Hz directly from the lock-in amplifier to the PC via an RS232 interface.

A 40-cm long fused silica capillary, 50-μm-ID, 190-μm-OD, was used for the separations after conditioning for at least 24 hours with running buffer. The same capillary was used for over five hundred runs and discarded only due to accidental breakage. The length from injection (anode) to detection was 35 cm. Separation buffer was 10.7 mM sodium phosphate, 1.8 mM sodium tetraborate and 25 mM SDS pH 6.7, which was filtered through a Millex-GS 0.22 μm syringe-filter unit before using. Samples were electrokinetically injected at 2 kV for 5 s and separated at 10 kV. Current through the capillary was approximately 8 μA.

The PTH amino acid standard (from ABI) was dissolved in acetonitrile during storage. A 5 μL aliquot was dried and redissolved in 5-10 μL of running buffer before injection. PTH residues from degradation were dissolved in 10 to 500 μL running buffer, depending on cycle yield, and electrokinetically injected into the capillary.

3.3 Results and Discussion

Edman Degradation

Insulin chain B from bovine and SP-5 (splenopentin) were the two peptides used in this study. The primary structure of these two samples is illustrated in Figure 3.1.

Insulin Chain B

$\mathrm{F - V - N - Q - \overset{5}{H} - L - C - G - S - \overset{10}{H} - L - V - E - A - \overset{15}{L} -}$

$\mathrm{Y - L - V - C - \overset{20}{G} - E - R - G - F - \overset{25}{F} - Y - T - P - K - \overset{30}{A}}$

SP-5, Splenopentin

$\mathrm{R - K - E - V - \overset{5}{Y}}$

Figure 3.1 Amino acid sequences of insulin chain B and SP-5, splenopentin.

In the Applied Biosystems automated peptide sequencer, 12.5% TMA buffer is delivered in the vapour phase providing the requisite pH 9.5 environment for coupling. However, the 12.5% TMA liquid solution has a pH of 12.5 so it was adjusted to pH 9.5 by addition of TFA. PITC is unstable at pH greater than 10 [13]. The coupling buffer also had to contain 60% propanol to facilitate co-dissolution of peptide and PITC. The order in which coupling buffer, peptide and PITC were added was important for the reaction. Presumably a uniform solution of these components is required. Coupling was initially done for 20 minutes at 55°C but yields were very low (25%). Coupling time was increased and temperature decreased to 30 minutes at 50°C producing at least a two-fold increase in yield.

After coupling, wash steps with heptane and ethyl acetate were discovered to be insufficient for removing organic impurities. Heptane was good at removing excess PITC but an additional wash with benzene was necessary to remove thiourea-based by-products. Ethyl acetate was necessary to remove residual TMA and water. TMA can react with TFA to form a salt and residual water can aid in acid hydrolysis of the peptide during cleavage.

Cleavage was initially done for 8 minutes at 55°C, but later changed to 15 minutes for 50°C because of appearance of PTC-peptide during PTH residue determination. Extraction of the ATZ amino acid product was originally done with chlorobutane, as in the commercial system. A large unidentified peak, eluting near PTH-arginine, was seen during the CE separations, and was attributed to a compound formed during the cleavage step. When the extraction solvent was switched to benzene, the amount of this unknown product was reduced but, as seen in some of the electropherograms, it was still present. Even when a blank degradation was performed (no peptide added), this by-product peak was seen. Boehnert and Schlesinger [45] indicate that repetitive yield is better with benzene than chlorobutane when manually sequencing peptides. Great care was taken when handling benzene because of its carcinogenic nature. In general, low yields (<50 %) were achieved for the manual sequencing of insulin and SP-5. These low yields are mainly attributed to the unconventional chemistry used, and lack of further optimization of reaction conditions.

Separation of PTH amino acid standard

In order to identify PTH amino acid products from protein sequencing, it is necessary to completely separate all PTH amino acids and by-products generated from Edman degradation. The major by-products of sequencing with phenylisothiocyanate (PITC) are diphenylthiourea (DPTU) and dimethylphenylthiourea (DMPTU). During the coupling step, PITC can undergo hydrolysis to form aniline. Reaction of aniline with a second PITC molecule forms DPTU. Also, PITC can react with residual dimethylamine, a contaminant in the TMA coupling buffer, to form DMPTU. Reactions for the formation of these two by-products are shown in Figure 3.2.

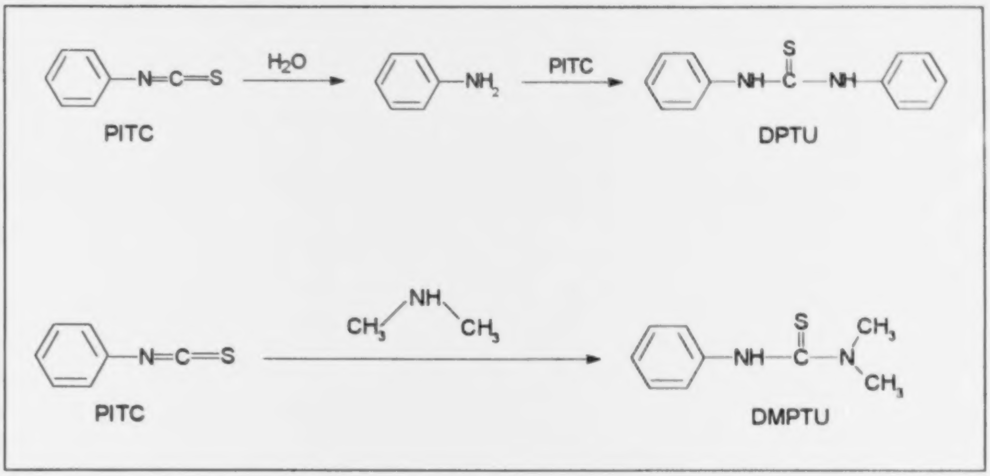


Figure 3.2 By-products formed during the coupling step of Edman degradation.

Typically, the PTH amino acids are identified by HPLC with UV detection in commercial protein sequencers. CE is not routinely used because the small volume requirements have limited its use with conventional absorbance detectors. The laser-based thermo-optical absorbance method, on the other hand, demonstrates femtomole limits of detection (Chapter 2) for PTH amino acids, which is about 1,000 times better than currently used methods for HPLC.

Most of the PTH amino acids and Edman by-products are neutral and cannot be separated by free zone capillary electrophoresis. Otsuka and coworkers [46] demonstrated the separation of PTH amino acids at picomole levels by micellar electrokinetic chromatography using SDS as an additive; however the quality of separation was not ideal for protein sequencing. Since the major limit of primary sequence determination by PITC degradation seems to be the HPLC analysis, improvements at this end are needed to accompany the recent improvements in sequencer design [14-16].

Sodium dodecyl sulfate (SDS) has proven to be an excellent surfactant for micelle formation. As Terabe and others [46-52] have demonstrated, neutral species can be partitioned within the micelles to allow their separation. Since the partition capacity of each PTH amino acid is dependent on the concentration of micelles, the total amount of SDS will effect the resolution of the separation. As the concentration of SDS increases above 8 mM (critical micelle concentration, CMC), the number of micelles in solution increases. The effect of SDS concentration in pH 7 buffer on PTH amino acid migration time is shown in Figure 3.3. Significantly, the 19 PTH amino acids and two by-products cannot all be totally separated with an arbitrary concentration of SDS above the CMC. However, if PTH-histidine is not considered, the other 18 PTH-amino acids can be completely separated in the 25 to 30 mM range for SDS. Since SDS weakly absorbs at 248 nm, the detection wavelength, a higher concentration of SDS gave a higher background signal. Therefore, 25 mM SDS was chosen as the optimum concentration.

As in HPLC determination [13], PTH-histidine is highly sensitive to the local environment during analysis with CE techniques. In this study the migration time of PTH-histidine shifted with changes in buffer composition, running voltage, and room temperature. This sensitive nature of PTH-histidine migration time arises from having a pKa near 6, similar to unlabeled histidine. Therefore, the retention time of PTH-histidine is highly dependent on pH of the running buffer, which is preferably kept near physiological pH. The migration time of each PTH-amino acid as a function of pH is shown in Figure 3.4. The curves in Figure 3.4 show that PTH-histidine migration time

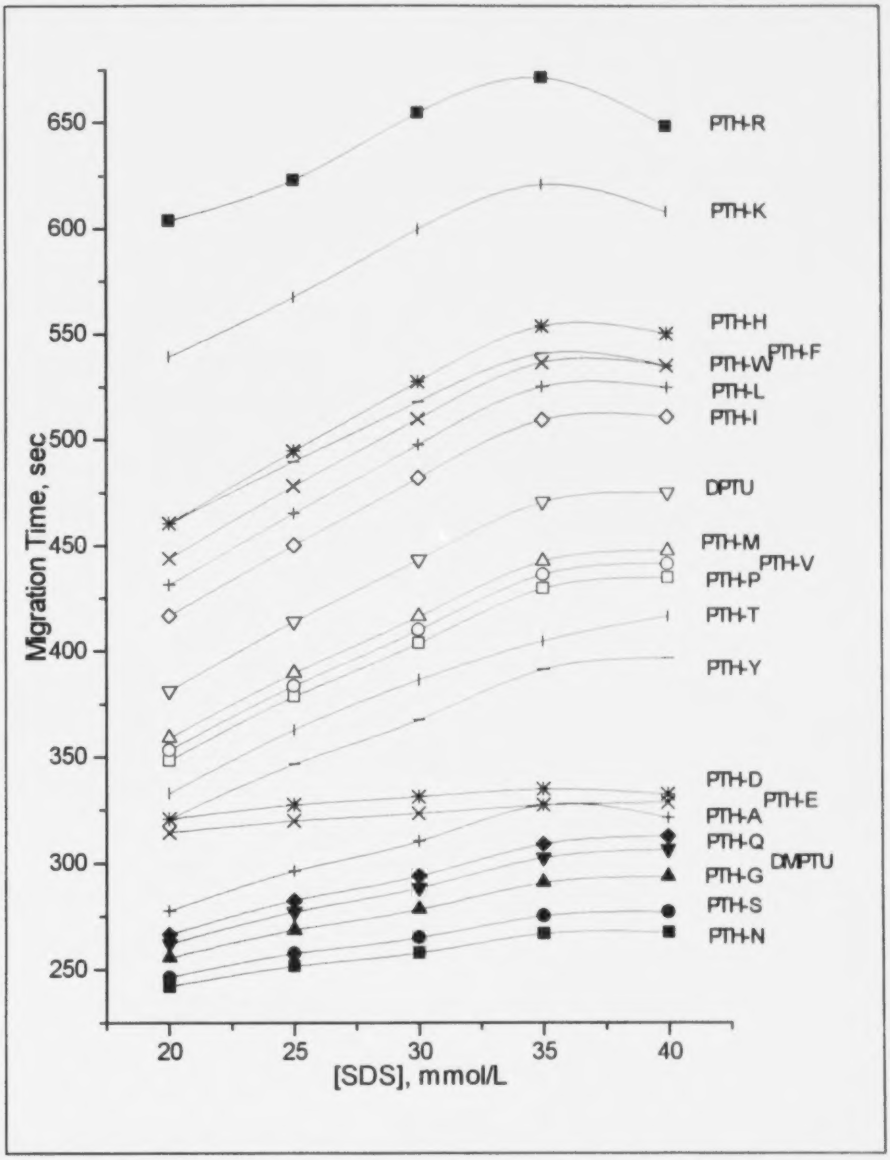


Figure 3.3 Effect of SDS concentration on migration time of the PTH amino acids. A 50-μm-ID, 40-cm-long capillary was used for separation: the distance from injector to detector was 35 cm. Injection was for 5s at 2 kV. Separation proceeded at 10 kV in a running buffer of 10 mM sodium phosphate, 2.5 mM sodium tetraborate, pH 7.0 with varying concentrations of SDS.

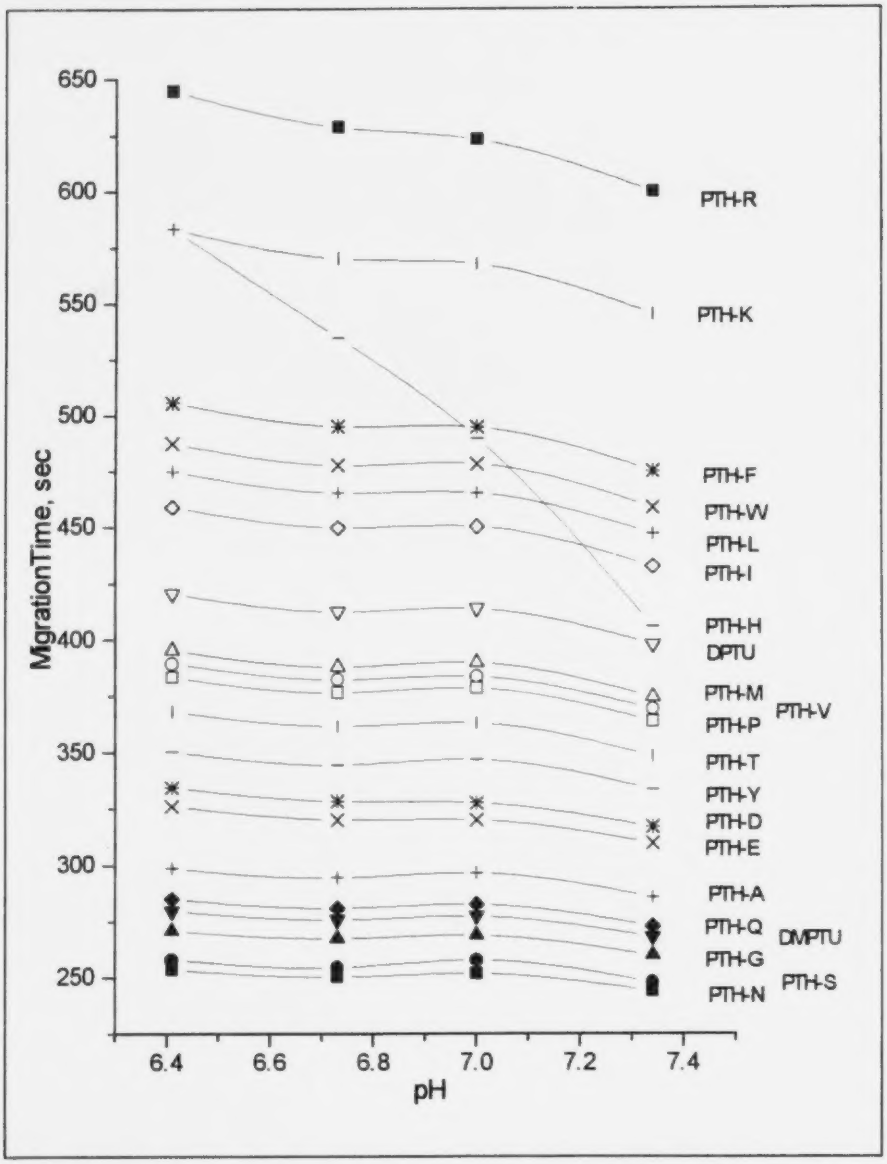


Figure 3.4 Effect of pH on the migration time of PTH amino acids. A 50-µm-ID, 40-cm-long capillary was used for separation: the distance from injector to detector was 35 cm. Injection was for 5s at 2 kV. Separations proceeded at 10 kV in a running buffer of 12.5 mM sodium phosphate/sodium tetraborate and 25 mM SDS. The sodium tetraborate concentration was incrementally increased, with concomitant decrease in sodium phosphate, to vary pH of the buffer.

varies by more than two minutes over a pH change of 1. The migration times of the other eighteen PTH-amino acids are essentially constant over a change of 1 pH unit. The optimized pH was found to be 6.7, with PTH-histidine migrating reproducibly between PTH-phenylalanine and PTH-lysine. Once buffer composition and pH were optimized, the order of elution of all peaks was always the same.

Migration time reproducibility is less than 1% RSD when capillaries are equilibrated with running buffer at least 24 hours before first use. Dilute hydroxide solutions are never used to rinse capillaries between runs, as suggested by some authors [52]. Rinsing with extremely high (or low) pH solutions alters the zeta potential established by the buffer's own ionic strength, composition and pH. The electropherogram of nineteen PTHs and two by-products determined by MECC with thermo-optical detection is shown in Figure 3.5. All species are completely separated. Besides the improved detection limit of PTH amino acids (Chapter 2), CE analysis provides a better separation than state-of-the-art RP-HPLC, and the CE separation is three times faster. To aid in peak identification of PTHs produced from the peptide degradation, a simple program was written to normalize migration times to that of DPTU, one of the major by-products of the Edman degradation.

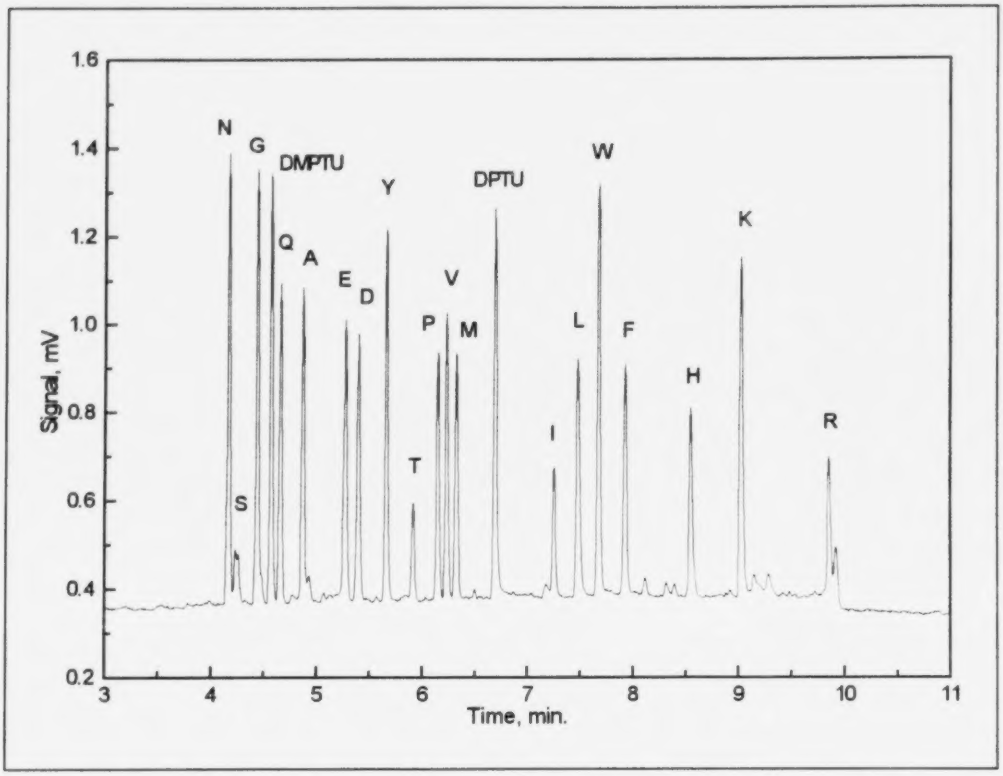


Figure 3.5 Capillary electropherogram of 60 fmol PTH amino acid standard containing nineteen PTHs, DPTU and DMPTU as purchased from ABI. Each PTH residue is designated by the single letter abbreviation of the corresponding amino acid. A 50-μm-ID 40-cm-long capillary was used for the separation: the distance from injector to detector was 35 cm. Injection was for 5 s at 2 kV. Separation proceeded at 10 kV in a 10.7 mM phosphate and 1.8 mM borate and 25 mM SDS pH 6.7 buffer.

Typical CE injection volumes are 1 nL. This volume constitutes, at most, 0.1% of the sample volume because at least 1 μL is needed to immerse both the capillary and the electrode into solution for an electrokinetic injection. Therefore, to inject one femtomole, about one picomole of PTH amino acid is necessary within the 1 μL sample volume. One picomole per microlitre equals 1×10^{-6} M, close to the concentration detection limit of HPLC. Obviously, a method of sample concentration would be preferred. Researchers have found several successful methods of sample concentration: isotachophoresis (ITP) [53], isoelectric focusing (IEF) [54], reversed-phase packing [55], and capillary bundles [56]. The use of reversed-phase packing seemed to be a logical method of in-column sample concentration since it is currently the separation method of choice (RP-HPLC) for PTH amino acids.

Preliminary studies showed that PTH amino acids are desorbed from C-18 (octadecylsilane) packing material by using 20 to 25 % acetonitrile in water. The sample could, therefore, be injected from 100% aqueous buffer, concentrated on a small plug of C-18 packing contained within the capillary, and eluted off in 20% acetonitrile buffer. Separation by MECC could occur once all PTHs were desorbed. The first step was to optimize the separation in an acetonitrile buffer system. With this scheme in mind, a separation of twenty PTH amino acids was attempted in a mixed organic/aqueous buffer. Figures 3.6A and 3.6B, show the electropherograms of 20 PTH amino acids separated in a 20% acetonitrile buffer with 35 and 50 mM SDS. Plates A and B of Figure 3.6 indicate that the effect of organic modifier is to reduce micelle formation. This effect is expected since SDS forms only normal (aqueous) solution micelles [57].

There is a problem, unfortunately, with concentrating PTHs on C-18 packing that makes the method infeasible for MECC separation. Monomers of SDS in the buffer preferentially adsorb to the packing material, rendering it useless for retaining PTH amino acids [58]. With this in mind, further study on MECC separations in organic modified buffers was halted before optimization of the separation in an acetonitrile/water buffer was complete. Rather than pursue methods for trying to concentrate the sample, it is easier to test the ability of CE/thermo-optical absorbance to analyse real sequencing samples by degrading larger amounts of peptides. The results of manually degrading two peptides are presented in the following section.

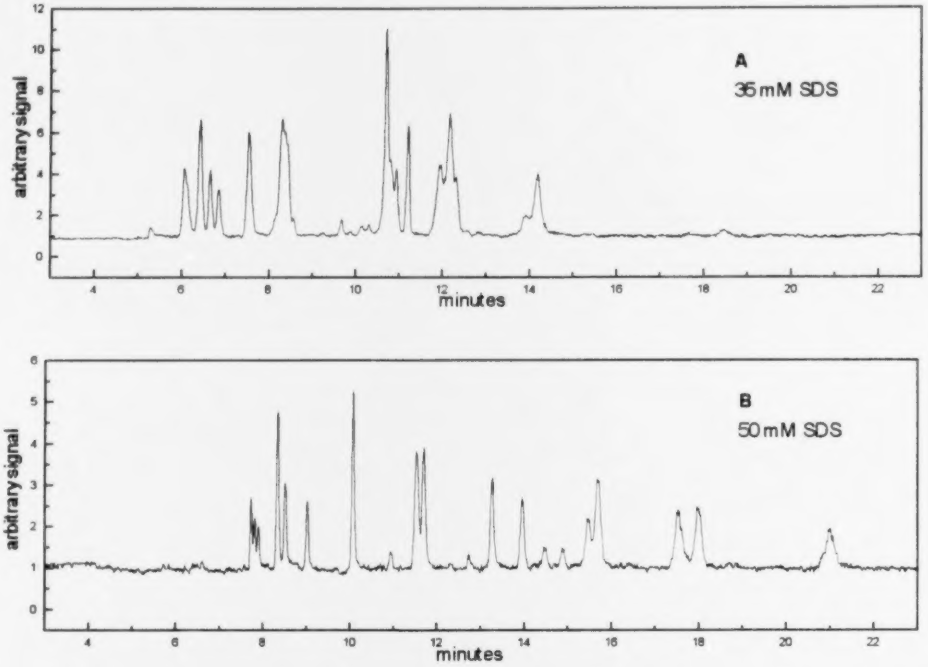


Figure 3.6 Electropherogram of twenty PTH amino acids in organic/aqueous buffer. A 50-µm-ID, 42-cm-long capillary was used for the separation: the distance from injector to detector was 37 cm. Sample was injected hydrodynamically for 10 s at a height of 2.5 cm. Separation proceeded at 10 kV in a 12.5 mM sodium phosphate/sodium tetraborate, 35 mM SDS (A), or 50 mM SDS (B), pH 7 buffer made up in 20 % acetonitrile. Approximately 65 fmol of each amino acid was injected.

Separation of Edman degradation products

The small peptide used in this study is SP-5, also called splenopentin, a five-residue molecule with the sequence: NH_2-arginine-lysine-glutamic acid-valine-tyrosine-COOH (NH_2-R-K-E-V-Y-COOH). Figure 3.7 shows the electropherograms for the sequence analysis of 865 nmol of SP-5. The standard represents approximately 20 fmol for each PTH amino acid, DPTU and DMPTU. No PTH-cysteine is present in the standard. The time required for stepwise degradation for each cycle is approximately 5 hours because of the long drying times used. After each cycle, the PTH residue was dissolved in 200 μL electrophoresis running buffer, determined, and re-diluted as necessary to give a reasonable peak height. Peaks must be within the linear range of peak height versus concentration to be able to estimate amounts. Each trace in Figure 3.7 represents a different dilution factor so it is difficult to see how much PTH amino acid each peak represents. The exact amounts determined are presented in Table 3.1.

Table 3.1 Quantitation of the sequence analysis of SP-5 peptide.

Cycle Number	Amount injected into CE system	PTH recovered from sequencing	Yield (%)
Cycle 1, PTH-arg	40 fmol	160 nmol	18
Cycle 2, PTH-lys	109 fmol	568 nmol	66
Cycle 3, PTH-glu	266 fmol	1360 nmol (?)	100 (?)
Cycle 4, PTH-val	59 fmol	715 nmol	83
Cycle 5, PTH- tyr	75 fmol	500 nmol	58

The values in Table 3.1 were calculated by comparing the peak height in each cycle to that in the standard, then compensating for the dilution and volume injected to obtain the amount of PTH recovered from sequencing per cycle. Repetitive yield (RY) for sequencing can be calculated by using the following equation [13]:

$$\mathrm{RY} = \left(\frac{Y_b}{Y_a}\right)^{\frac{1}{b-a}} \cdot 100, \quad a < b \tag{3.1}$$

where Y_a and Y_b are the yields at cycles a and b respectively. The repetitive yield for the results of cycle 2 and 5 as shown in Figure 3.7 is 95 %. There are some doubts as to how accurate 95 % is since the yields increased from cycle 1 to 5. Also, the result for cycle 3 is ridiculous and indicates that there may have been a mistake in the injection or the peak for PTH-E is in a non-linear region of the dynamic range for the separation/detection system. The electropherograms show good signal-to-noise for the analyte PTH residue because of the very large amount of starting material used. In the first cycle, a few unidentified peaks are seen besides the DPTU by-product peak. Extensive wash steps allow complete removal of DMPTU. Each cycle shows a slight amount of lag: PTH product from the previous cycle present in the current cycle. Only cycle 2 shows evidence of preview: PTH product from the following cycle present in the current cycle. Problems such as preview and lag are more pronounced in automated instruments because rigorous washing steps are more difficult to implement in pre-programed cycles.

Initially, attempts at sequencing SP-5 were not successful because low nanomole quantities were used and the truncated peptide would get lost in the wash steps. Increasing the starting amount of peptide was helpful and addition of Polybrene provided the best way for retaining the peptide from cycle to cycle.

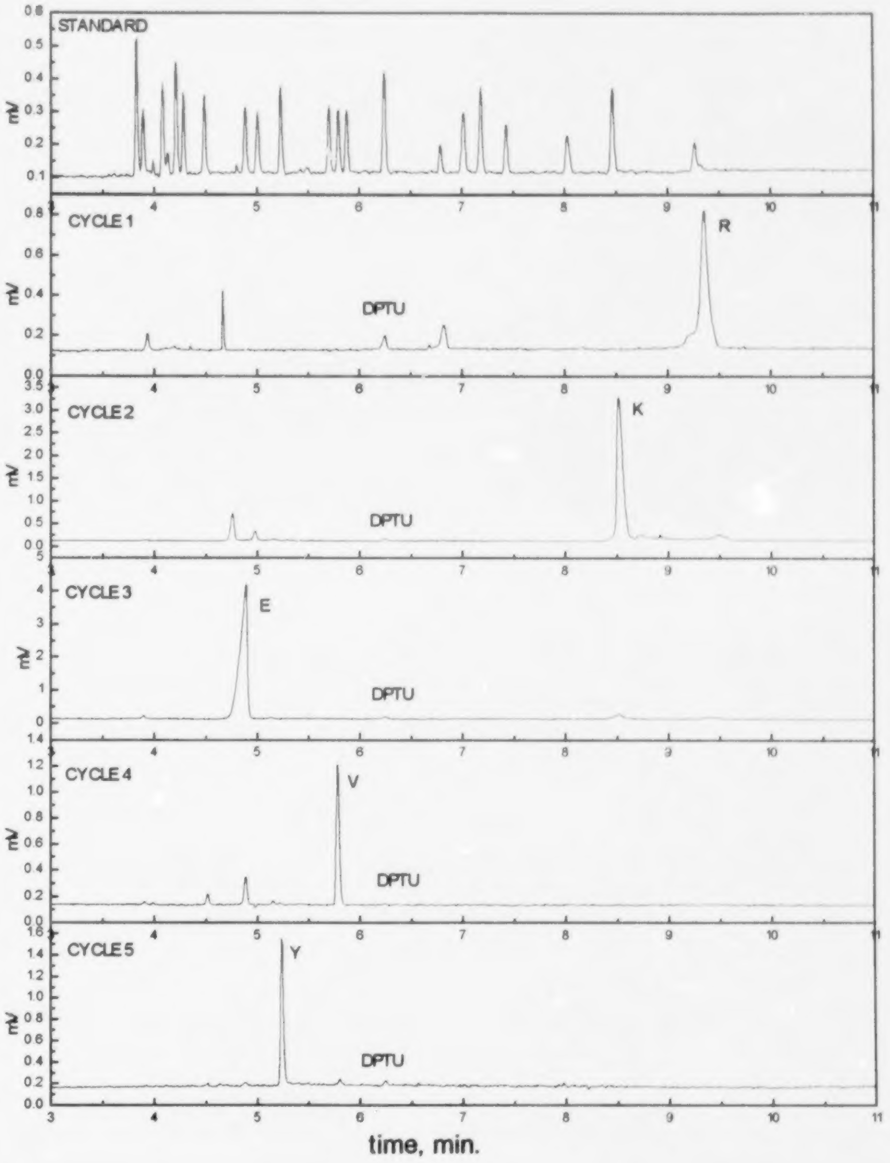


Figure 3.7 Electropherograms for the sequence analysis of SP-5 peptide. SP-5 (865 nmol) was sequenced by manual Edman degradation and analysed by MECC/thermo-optical absorbance detection as described in the text. Separation conditions are the same as for Figure 3.5. The by-product peak DPTU is used to normalize migration times to the standard (see Appendix C).

The second peptide used in this study was insulin chain B, 30 residues in length, the sequence for which was shown in Figure 3.1. The results from manual Edman degradation of 98 nmol of insulin chain B are presented in Figures 3.8 A to F. For each set of electropherograms, the standard run on the same or closest day to the PTH determination is shown at the top of the figure. Each peak in each standard represents approximately 20 fmol PTH amino acid, DPTU or DMPTU and is identified in Figure 3.5. Results from the 28th and 30th cycles are not presented because these samples were discarded by mistake during extraction and washing steps. As with the SP-5 peptide, dried PTH residues were dissolved in 20 to 100 µL of electrophoresis running buffer then approximately 0.3 nL injected. Table 3.2 lists the amount of degradation product determined and the amount obtained from sequencing, corrected for dilution. Different amounts were injected into the CE system for determination. PTH recovered is the total amount of PTH amino acid collected after the conversion step of the degradation for each cycle.

The repetitive yield for sequencing insulin chain B, calculated using Equation 3.1 and cycles 1 and 10, is 72%. Seven of twenty-eight cycles gave no sequence information and from the seventeenth cycle on, analyte signal-to-noise decreased rapidly as background peaks appeared more often. Lack of information at PTH-cysteine is typical, caused by β-elimination of H_2S. PTH-cysteine was not included in the standard because of this problem, a practice followed by many researchers [59]. Lag is seen in several cycles in Figure 3.8 and preview in a few instances. A more bothersome problem, of unknown origin, is the large peak seen near or after 10 minutes in many of the electropherograms. Attempts were made to identify this peak but it cannot be attributed to PTC-peptide, DPU or other obvious by-products. A similar phenomenon is seen in the chromatograms for sequence analysis of a protein in Calaycay *et al.* [15]. The peak near 10 minutes generally gets larger the further sequencing goes and was originally much larger when chlorobutane was being used to extract the ATZ amino acid after cleavage. Extraction with benzene seemed to reduce the intensity of this peak. Species that migrate slowly in an MECC separation are either cationic or very non-polar, having a large partition coefficient for the micellar phase. Since this anomalous peak only interferes with identification of PTH-arginine, its identity was not treated as a major concern.

Table 3.2 Quantitation of the sequence analysis of insulin chain B.

Cycle No. and residue	Amount injected (fmol)	PTH recov'd (nmol)	Yield (%)	Cycle No. and residue	Amount injected (fmol)	PTH recov'd (nmol)	Yield (%)
1, PTH-phe	84	32	39	15, PTH-leu	33	2.9	3
2, PTH-val	97	29	29	16, PTH-tyr	35	1.9	2
3, PTH-asn	--	--	--	17, PTH-leu	12	0.9	1
4, PTH-gln	205	42	43	18, PTH-val	13	1.0	1
5, PTH-his	34	6.8	7	19, PTH-cys	--	--	--
6, PTH-leu	27	8.3	8	20, PTH-gly	20	1.0	1
7, PTH-cys	--	--	--	21, PTH-glu	10	1.0	1
8, PTH-gly	148	7.4	8	22, PTH-arg	--	--	--
9, PTH-ser	43	2.2	2	23, PTH-gly	45	2.2	2
10, PTH-his	20	1.7	2	24, PTH-phe	65	5.0	5
11, PTH-leu	32	2.5	3	25, PTH-phe	7	0.6	1
12, PTH-val	77	4.7	5	26, PTH-tyr	--	--	--
13, PTH-glu	93	5.0	5	27, PTH-thr	--	--	--
14, PTH-ala	130	6.7	7	29, PTH-lys	--	--	--

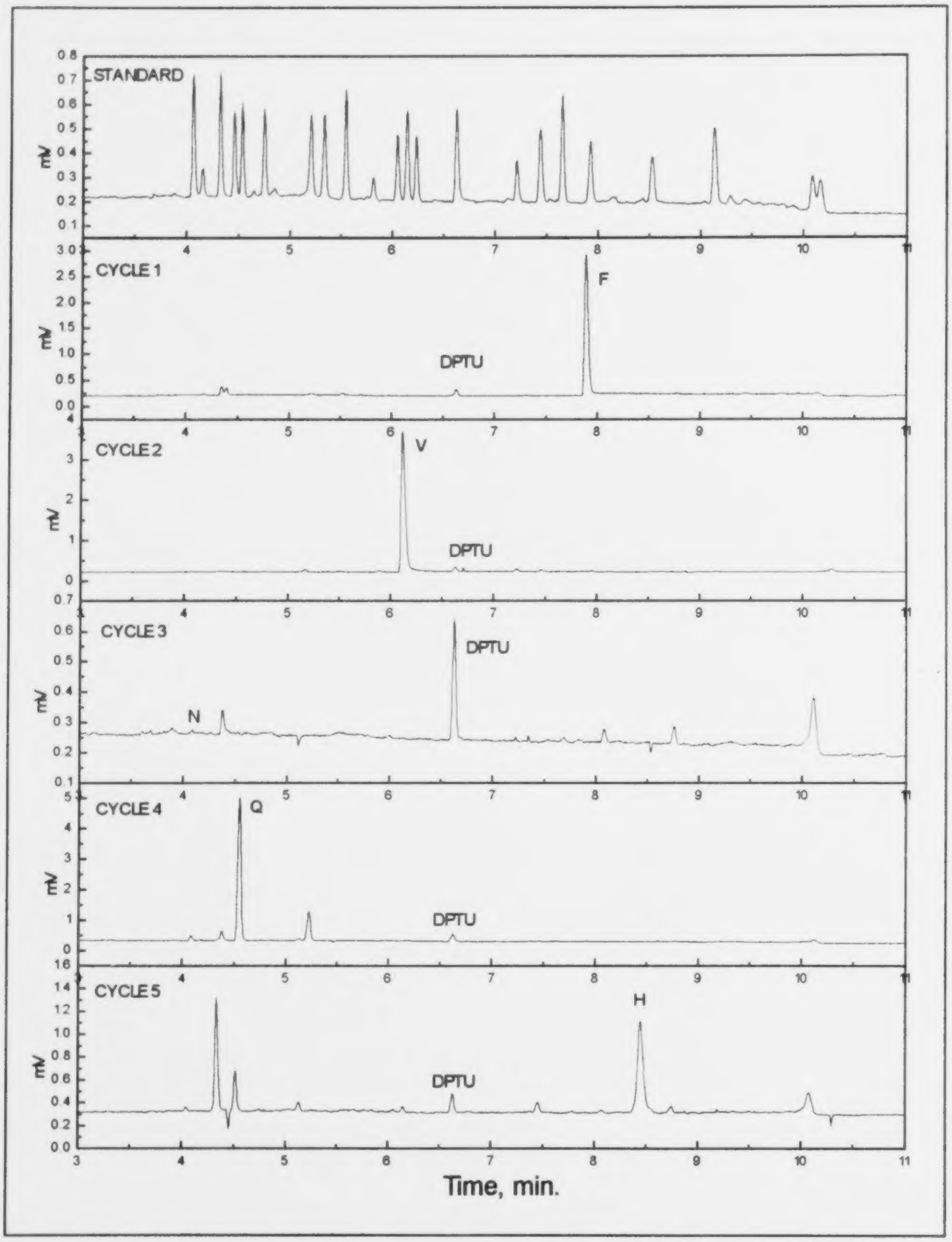


Figure 3.8A Electropherograms for the sequence analysis of insulin chain B, cycles 1 to 5. Insulin chain B (98 nmol) was sequenced by manual Edman degradation and analysed by MECC/thermo-optical absorbance detection as described in the text. Separation conditions are the same as for Figure 3.5. The by-product peak DPTU is used to normalize migration times to the standard (see Appendix C).

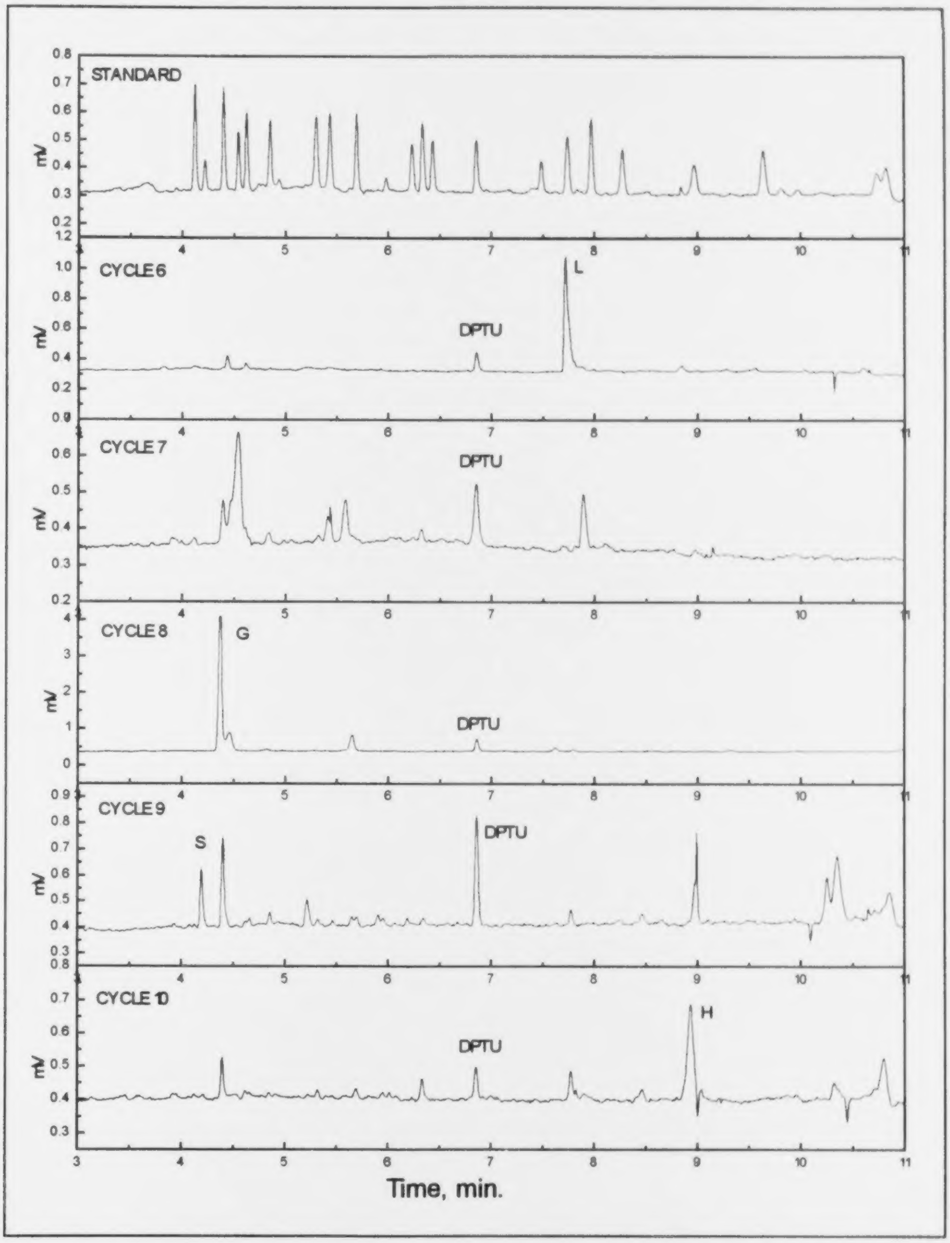


Figure 3.8B Electropherograms for the sequence analysis of insulin chain B, cycles 6 to 10. Insulin chain B (98 nmol) was sequenced by manual Edman degradation and analysed by MECC/thermo-optical absorbance detection as described in the text. Separation conditions are the same as for Figure 3.5. The by-product peak DPTU is used to normalize migration times to the standard (see Appendix C).

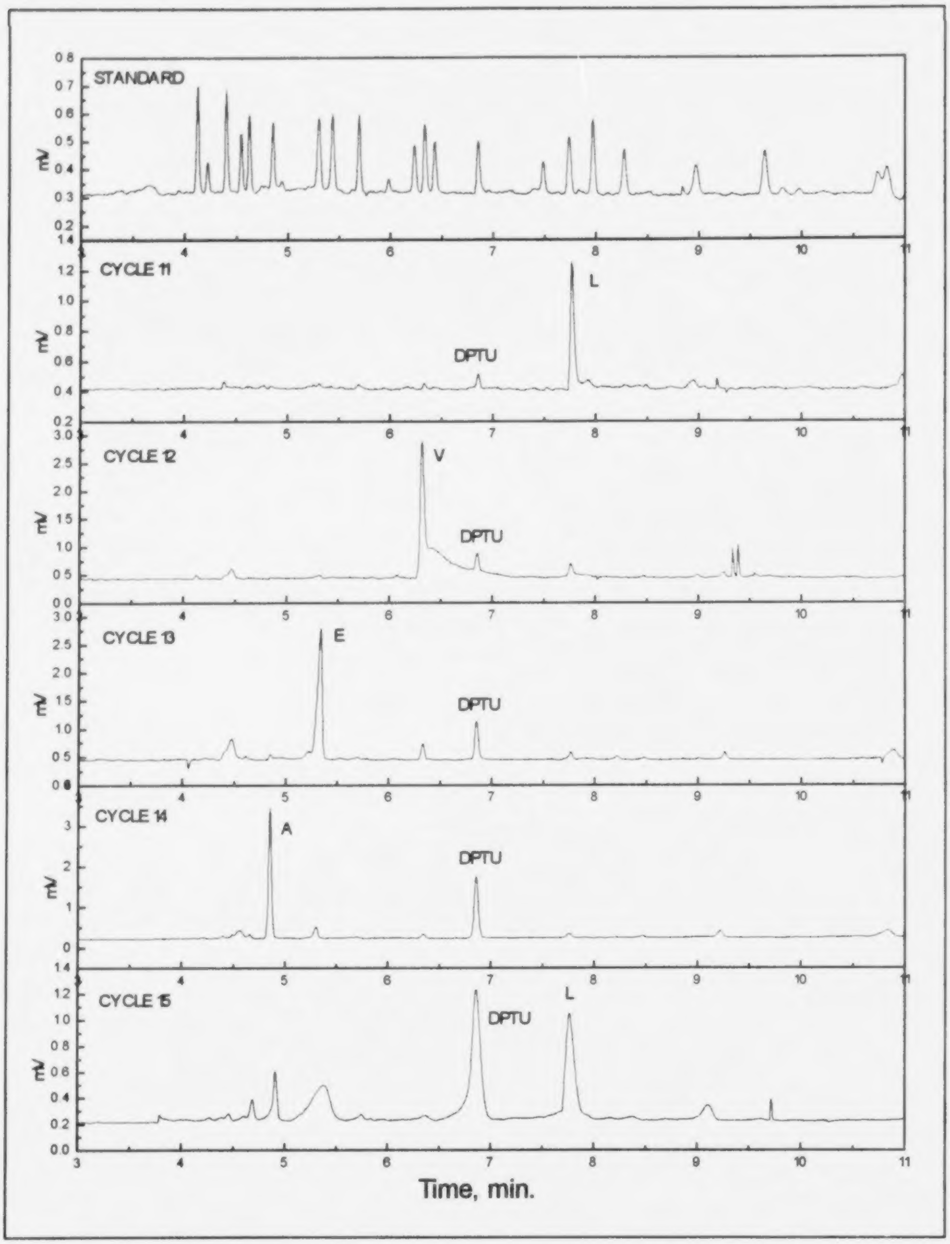


Figure 3.8C Electropherograms for the sequence analysis of insulin chain B, cycles 11 to 15. Insulin chain B (98 nmol) was sequenced by manual Edman degradation and analysed by MECC/thermo-optical absorbance detection as described in the text. Separation conditions are the same as for Figure 3.5. The by-product peak DPTU is used to normalize migration times to the standard (see Appendix C).

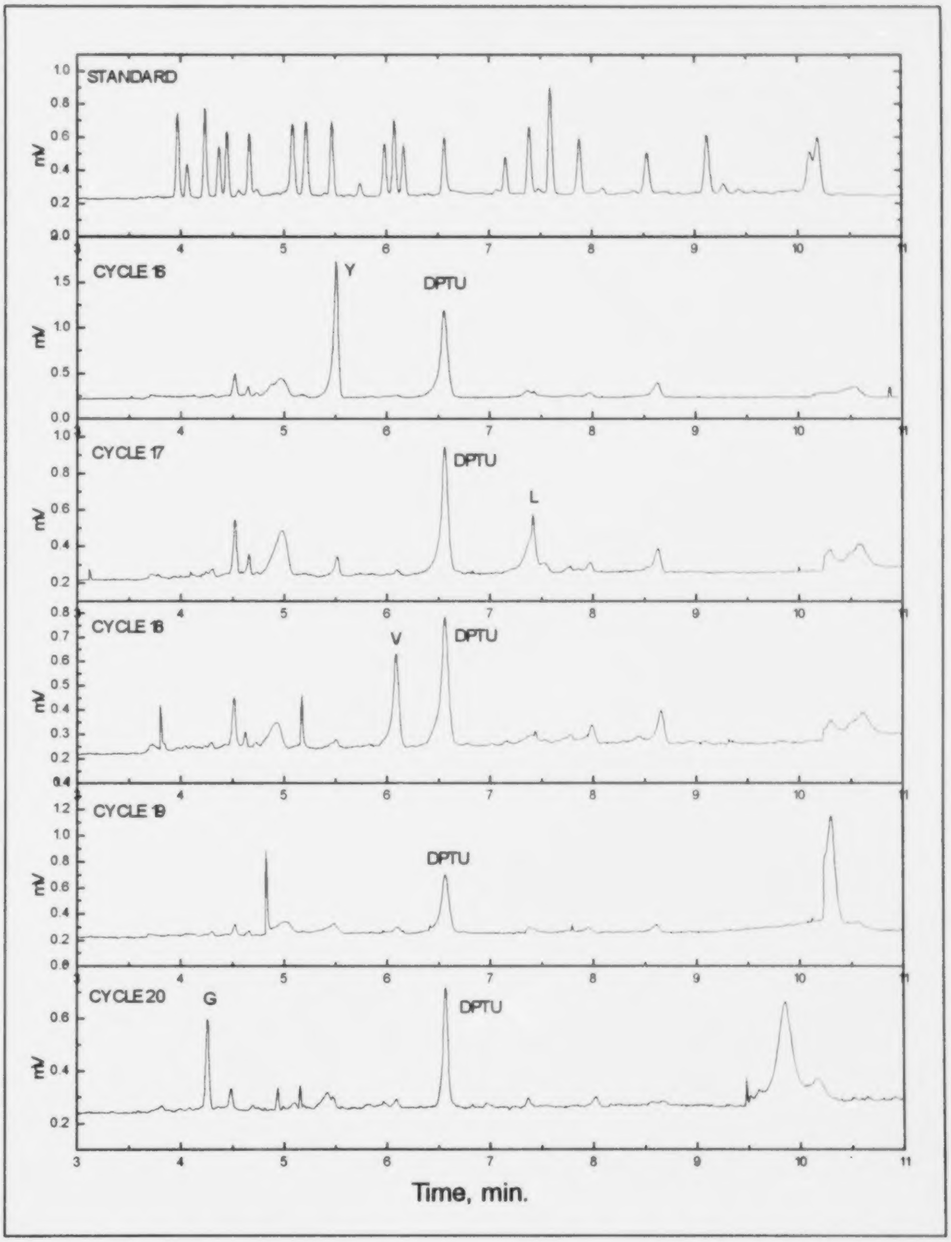


Figure 3.8D Electropherograms for the sequence analysis of insulin chain B, cycles 16 to 20. Insulin chain B (98 nmol) was sequenced by manual Edman degradation and analysed by MECC/thermo-optical absorbance detection as described in the text. Separation conditions are the same as for Figure 3.5. The by-product peak DPTU is used to normalize migration times to the standard (see Appendix C).

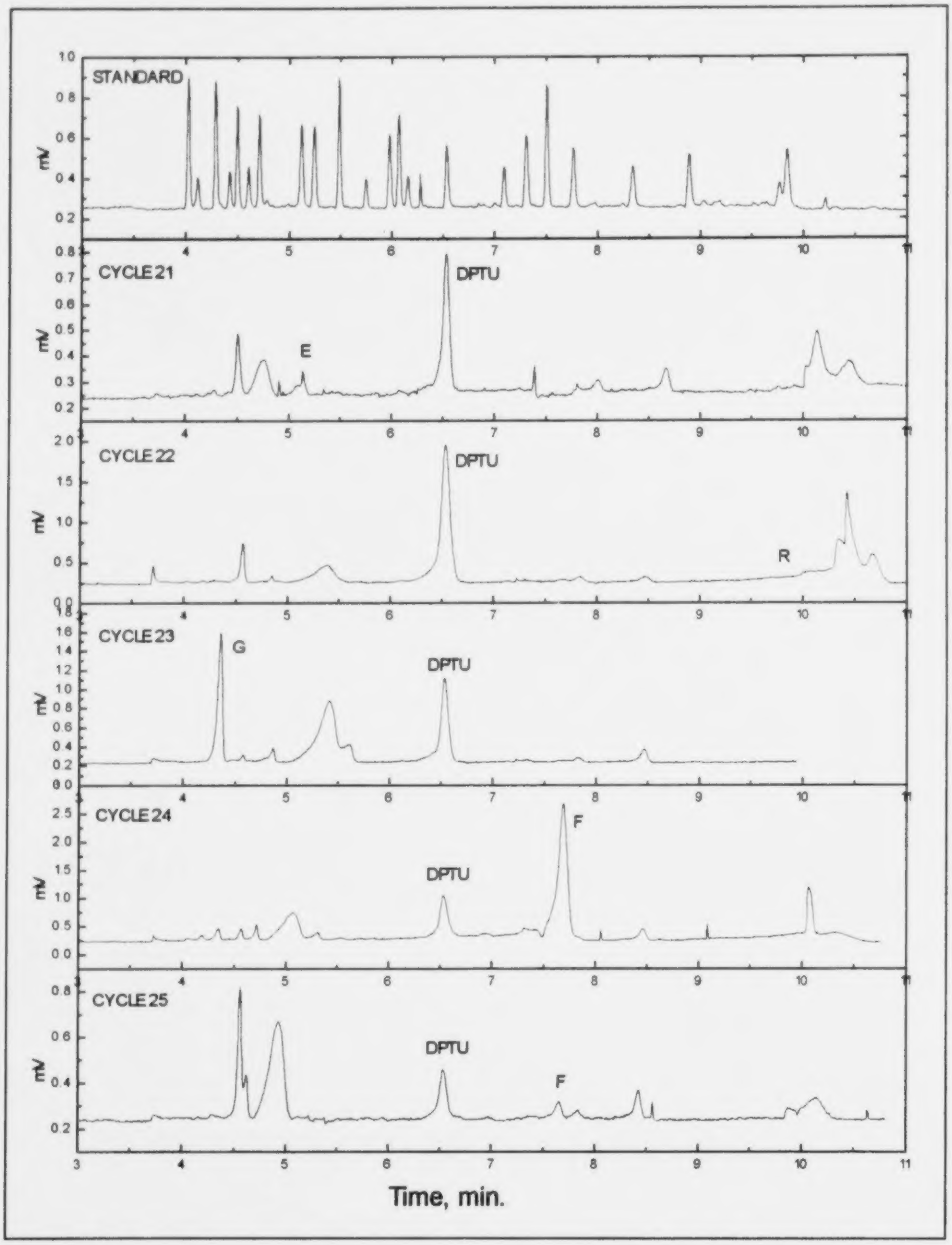


Figure 3.8E Electropherograms for the sequence analysis of insulin chain B, cycles 21 to 25. Insulin chain B (98 nmol) was sequenced by manual Edman degradation and analysed by MECC/thermo-optical absorbance detection as described in the text. Separation conditions are the same as for Figure 3.5. The by-product peak DPTU is used to normalize migration times to the standard (see Appendix C).

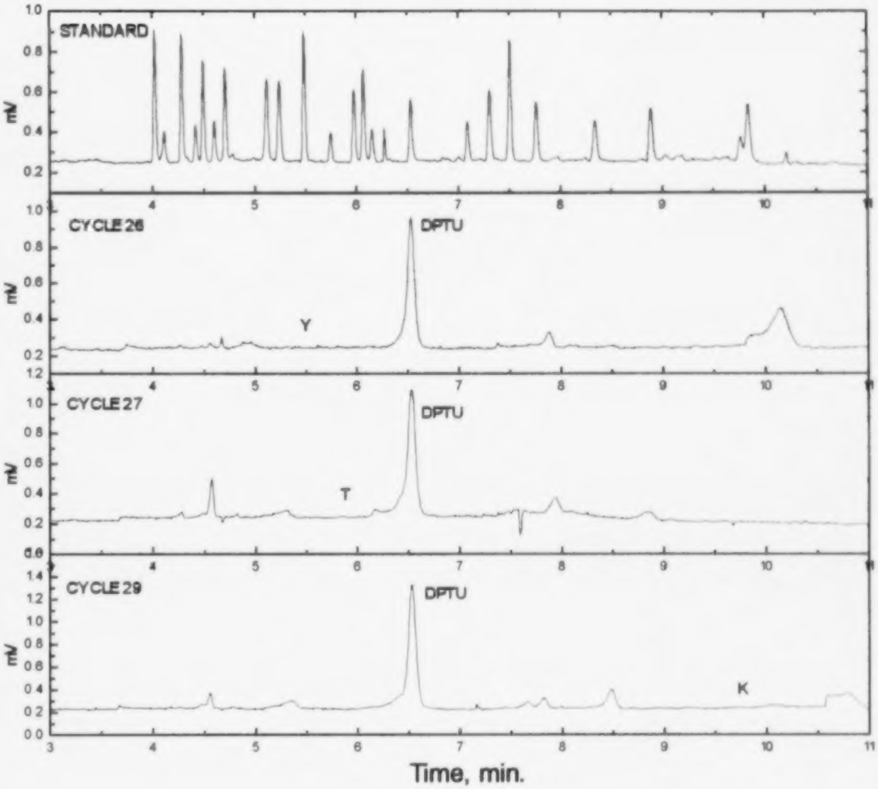


Figure 3.8F Electropherograms for the sequence analysis of insulin chain B, cycles 26, 27 and 29. Insulin chain B (98 nmol) was sequenced by manual Edman degradation and analysed by MECC/thermo-optical absorbance detection as described in the text. Separation conditions are the same as for Figure 3.5. The by-product peak DPTU is used to normalize migration times to the standard (see Appendix C).

The sequencing results for SP-5 and insulin chain B were not performed at the sub-picomole or even picomole level because of difficulties with the manual sequencing protocol. Even thirteen years ago, manual Edman sequencing down to one nanomole was reported [60]. Interestingly, these authors referred to their method as micro quantitative sequencing, in 1980. In 1988, Haniu and Shively [61] described manual gas-phase microsequence analysis using a small volume continuous flow reactor made from a Teflon tube. They reported that good sequencing results could be obtained in the 100-200 pmol range for peptides and proteins. PTH determination was by reversed-phase HPLC, the same as in commercial automated instruments. These results imply that manual sequencing methods are at least 100 times less sensitive than automated ones. Automated gas-phase and solid-phase sequencing instruments are needed for picomole sequencing.

Although the Edman degradation procedure was not optimized, the results shown in Figures 3.7 and 3.8 demonstrate that sub-picomole amounts of PTH amino acid can be determined. The MECC separation of 19 PTH amino acids is complete in 10 minutes, three times faster than most reversed-phase HPLC separations. Actually, one can consider the MECC separation to be about 5 times faster because the HPLC solvents must be pumped out to re-set the gradient for a second run. This procedure requires an extra 15 to 25 minutes after HPLC analysis is complete. With MECC, a second sample can be analysed immediately after the first. One of the many advantages of MECC and other capillary electrophoretic technique is that the baseline is flat and there are no baseline disturbances near the start of the separation like those seen in HPLC. Solvent use is minimal and inexpensive because only water has to be purified. Also, capillary electrophoresis equipment is less costly than HPLC equipment: no high pressure pumps are required for CE and one meter of capillary column costs about $3.00, about two orders of magnitude less than for a packed HPLC column. Good reproducibility in migration times is seen from the data presented in Figures 3.7 and 3.8. The elution order always stays the same, unlike in HPLC separation of PTH amino acids where aspartic acid and glutamic acid tend to move around because their retention times change as the column ages [62].

Thermo-optical absorbance detection provides superior mass detection limits compared to the UV absorbance detectors built into HPLCs. Most of the efforts to improve detection of PTH amino acids have revolved around trying to get more or all of the converted PTH into the injection loop by dissolving it in less solvent. Elaborate adaptions to the HPLC sample injector have been made in an attempt not to waste any product while still completely filling a 50 μL injection loop [15, 63]. HPLC technology is well established and few modifications have appeared in the literature recently. Perhaps reversed-phase HPLC cannot be improved and the only way to get better sensitivity for determination of PTHs is to explore new techniques. Capillary electrophoresis is emerging as a new method of analysis, and may be found to rival HPLC in the future, as the results in this chapter imply.

3.4 Conclusions

The amount of peptide used for the sequence analyses presented in this chapter did not demonstrate sub-picomole sequencing. However, the CE/thermo-optic determination of the PTH residues did demonstrate sub-picomole determination. This determination capability should translate into the ability to sequence femtomole amounts of polypeptides. The first section of Chapter 3 presented some background on current research being done to try to sequence less than a picomole of peptide or protein. One way to get more sequence information from less than 1 pmol is to improve detection of the sequencing products, the PTH amino acids. In the following sections of Chapter 3, MECC separation and thermo-optical absorbance detection was presented as a tool for identifying the products of manually degraded peptide. Optimization of the separation was shown to produce analyses three times faster than the current state-or-the-art in HPLC. Results of the sequencing and the residue identification were discussed in section 3.3. The manual Edman method described here was not sufficient for picomole and sub-picomole sequencing. Improvements in the chemistry would be necessary to sequence smaller peptide samples but the aim of this chapter was not to improve the Edman degradation, only the product identification. The MECC separation and thermo-optical detection of the PTH amino acid products showed fast, sensitive determination of the Edman sequencing products.

The reason CE has not previously been coupled to a commercially available sequencer is the incompatibility of volume requirements. Commercial sequencing instruments produce at least 50 µL of PTH amino acid solution whereas injection volumes in CE are on the order of one nanolitre. CE with thermo-optical absorbance detection has superb mass detection limits (<1 fmol) but only moderate concentration detection limits (10^{-6} M). On-column concentration methods for CE are available [53] but not at the 50-µL scale. Given these limitations, development of a capillary sized sequencing instrument, which has on-line injection of the PTH residues into the CE system, would add significantly to the ability of chemists to sequence trace amounts of polypeptides.

3.5 References

1. R.M. Hewick, M.W. Hunkapiller, L.E. Hood and W.J. Dreyer, *J. Biol. Chem.* **256**, 7990-7997 (1981).
2. A. Aitken, M.J. Geisow, J.B.C. Findlay, C. Holmes and A. Yarwood in *Protein Sequencing, a Practical Approach,* (eds. J.B.C. Findlay & M.J. Geisow) 43-68 (IRL Press; Oxford, 1989).

3. K.L. Stone and K.R. Williams in *A Practical Guide to Protein and Peptide Purification for Microsequencing,* (eds. P. Matsudaira) 43-69 (Academic Press, Inc.; San Diego, CA, 1993).

4. C.T. Mant and R.S. Hodges, ed., *High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis, and Conformation.* (CRC Press; Boca Raton, 1991).

5. R.H. Aebersold, D.B. Teplow, L.E. Hood and S.B.H. Kent, *J. Biol. Chem.* **261**, 4229-4238 (1986).

6. T. Choli and B. Wittmann-Liebold, *Electrophoresis* **11**, 562-568 (1990).

7. T. Bergman and H. Jornvall, *Electrophoresis* **11**, 569-572 (1990).

8. P. Matsudaira, *J. Biol. Chem.* **262**, 10035 (1987).

9. S.D. Patterson, D. Hess, T. Yungwirth and R. Aebersold, *Anal. Biochem.* **202**, 193-203 (1992).

10. P.W.M. Reisinger, T. Kleinschmidt and U. Welsch, *Electrophoresis* **13**, 65-72 (1992).

11. D.F. Reim and D.W. Speicher, *Anal. Biochem.* **207**, (1992).

12. J. Mozdzanowski and D.W. Speicher, *Anal. Biochem.* **207**, 11-18 (1992).

13. L.B. Smillie and M.R. Carpenter in *High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis and Conformation,* (eds. C.T. Mant & R.S. Hodges) 875-894 (CRC Press; Boca Raton, 1991).

14. S. Kent, L. Hood, R. Aebersold, D. Teplow, L. Smith, V. Farnsworth, P. Cartier, W. Hines, P. Hughes and C. Dodd, *BioTechniques* **5**, 314-321 (1987).

15. J. Calaycay, M. Rusnak and J.E. Shively, *Anal. Biochem.* **192**, 23-31 (1991).

16. N.F. Totty, M.D. Waterfield and J.J. Hsuan, *Protein Science* **1**, 1215-1224 (1992).

17. H.A. Scoble, J.E. Vath, W. Yu and S.A. Martin in *A Practical Guide to Protein and Peptide Purification for Microsequencing,* (eds. P. Matsudaira) 125-153 (Academic Press, Inc.; San Diego, CA, 1993).

18. H. Hirano, S. Komatsu, H. Takakura, F. Sakiyama and S. Tsunasawa, *J. Biochem.* **111**, 754-757 (1992).

19. R.G. Krishna, C.C.Q. Chin and F. Wold, *Anal. Biochem.* **199**, 45-50 (1991).

20. R. Aebersold, G.D. Pipes, R.E. Wettenhall, H. Nika and L.E. Hood, *Anal. Biochem.* **187**, 56-65 (1990).

21. J.M. Coull, D.J. Pappin, J. Mark, R. Aebersold and H. Koster, *Anal. Biochem.* **194**, 110-120 (1991).

22. S.-P. Liang, T.T. Lee and R.A. Laursen, *Anal. Biochem.* **197**, 163-167 (1991).

23. A. Iwamatsu, *Electrophoresis* **13**, 142-147 (1992).

24. D. Atherton, J. Fernandez and S.M. Mische, *Anal. Biochem.* **212**, 98-105 (1993).

25. D.C. Brune, *Anal. Biochem.* **207**, 285-290 (1992).

26. H.C. Krutzsch and J.K. Inman, *Anal. Biochem.* **209**, 109-116 (1993).

27. R.E. Wettenhall, R.H. Aebersold and L.E. Hood, *Methods in Enzymology.* **201**, 186-199 (1991).

28. J.Y. Chang, E.H. Creaser and K.W. Bentley, *Biochem. J.* **153**, 607-611 (1976).

29. H. Maeda, N. Ishida, H. Kawauchi and K. Tuzimura, *J. Biochem.* **65**, 777-783 (1969).

30. K. Muramoto, H. Kawauchi and K. Tuzimura, *Agric. Biol. Chem.* **42**, 1559-1563 (1978).

31. K. Muramoto, H. Kamiya and H. Kawauchi, *Anal. Biochem.* **141**, 446-450 (1984).

32. H. Kawauchi, K. Muramoto and J. Ramachandran, *Int. J. Peptide Protein Res.* **12**, 318-324 (1978).

33. A.S. Bhown, J.E. Mole, W.L. Hollaway and J.C. Bennett, *J. Chromatogr.* **156**, 35-41 (1978).

34. B. Wittmann-Liebold, J. Shan-Wei and J. Salnikow in *Protein/Peptide Sequence Analysis: Current Methodologies,* (eds. A.S. Bhown) 119-134 (CRC Press; Boca Raton, 1988).

35. R. Aebersold, E.J. Bures, M. Namchuck, M.H. Goghari, B. Shushan and T.C. Covey, *Protein Science.* **1**, 494-503 (1992).

36. K.C. Waldron, S. Wu, C.W. Earle, H.R. Harke and N.J. Dovichi, *Electrophoresis* **11**, 777-780 (1990).

37. S. Wu and N.J. Dovichi, *Talanta* **39**, 173-178 (1992).

38. Z. Deyl and R. Struzinsky, *J. Chromatogr.* **569**, 63-122 (1991).

39. M.A. Strege and A.L. Lagu, *Anal. Biochem.* **210**, 402-410 (1993).

40. M. Albin, S.-M. Chen, A. Louie, C. Pairaud, J. Colburn and J. Wiktorowicz, *Anal. Biochem.* **206**, 382-388 (1992).

41. J.K. Towns and F.E. Regnier, *J. Chromatogr.* **516**, 69-78 (1990).

42. H.H. Lauer and D. McManigill, *Anal. Chem.* **58**, 166-170 (1986).

43. J.T.K. Pang, D.J. Kramer and T.R. Tullsen in *Current research in protein chemistry: techniques, structure and function,* (eds. J.J. Villafranca) 11-22 (Academic Press, Inc.; San Diego, CA, 1990).

44. K.C. Waldron and N.J. Dovichi, *Anal. Chem.* **64**, 1396-1399 (1992).

45. M. Boehnert and D.H. Schlesinger, *Anal. Biochem.* **96**, 469-473 (1979).

46. K. Otsuka, S. Terabe and T. Ando, *J. Chromatogr.* **332**, 219-226 (1985).

47. S. Terabe, K. Otsuka, K. Ichikawa, A. Tsuchiya and T. Ando, *Anal. Chem.* **56**, 111-113 (1984).

48. S. Terabe, K. Otsuka and T. Ando, *Anal. Chem.* **57**, 834-841 (1985).

49. K. Otsuka, S. Terabe and T. Ando, *J. Chromatogr.* **348**, 39-47 (1985).

50. E.L. Little and J.P. Foley, *J. Microcol. Sep.* **4**, 145-154 (1992).

51. A.S. Cohen, S. Terabe, J.A. Smith and B.L. Karger, *Anal. Chem.* **59**, 1021-1027 (1987).

52. J. Liu, K.A. Cobb and M. Novotny, *J. Chromatogr.* **468**, 55-65 (1988).

53. R.-L. Chien and D.S. Burgi, *Anal. Chem.* **64**, 489A-496A (1992).

54. R. Aebersold and H. Morrison, *J. Chromatogr.* **516**, 79-88 (1990).

55. T. Tsuda in *Capillary Electrophoresis Technology,* (eds. N.A. Guzman) 489-508 (Marcel Dekker, Inc.; New York, 1993).

56. N.A. Guzman and M.A. Trebilcock, *Analitica Chimica Acta* **249**, 247-255 (1991).

57. L.J. Cline-Love, J.G. Habarta and J.G. Dorsey, *Anal. Chem.* **56**, 1132A-1148A (1984).

58. M. Chen, *Personal Communication* (1992).

59. M.R. Carpenter, *Personal Communication* (1992).

60. C. Zalut, W.J. Henzel and H.W. Harris, *J. Biochem. Biophys. Meth.* 11-31 (1980).

61. M. Haniu and J.E. Shively, *Anal. Biochem.* **173**, 296-306 (1988).

62. T. Bures, *Personal Communication* (1992).

63. G.S. Begg and R.J. Simpson in *Methods in Protein Sequence Analysis, Proceedings of the 7th International Conference, Berlin, July 3-8, 1988,* (eds. B. Wittmann-Liebold) 108-111 (Springer-Verlag; Berlin, 1989).

Appendix C: Normalization of migration times.

The standard mixture of PTH amino acids and the PTH amino acid collected after degradation each contain diphenylthiourea (DPTU). Therefore, DPTU can be used like an internal standard.

Assumptions:

(a) the relative mobility of DPTU to PTH is constant,

(b) variation in migration time between electropherograms is caused by electroosmotic flow differences,

(c) the electrophoretic mobility of all species is constant.

Equations:

(1) $\mu_{\mathrm{DPTU}} - \mu_{\mathrm{eo}} = \dfrac{\mathrm{L}}{\mathrm{E} \cdot t_{\mathrm{DPTU}}}$ where μ_{DPTU} is the electrophoretic coefficient of mobility, μ_{eo} is the electroosmotic coefficient of mobility, L is the capillary length, E is the electric field strength and t_{DPTU} is the migration time of DPTU in the run.

(2) $\mu_{\mathrm{PTH}} - \mu_{\mathrm{eo}} = \dfrac{\mathrm{L}}{\mathrm{E} \cdot t_{\mathrm{PTH}}}$ where μ_{PTH} is the electrophoretic coefficient of mobility and t_{PTH} is the migration time of PTH amino acid in the run.

(3) subtract (2)-(1), $\mu_{\mathrm{PTH}} - \mu_{\mathrm{DPTU}} = \dfrac{\mathrm{L}}{\mathrm{E}}\left(\dfrac{1}{t_{\mathrm{PTH}}} - \dfrac{1}{t_{\mathrm{DPTU}}}\right) = \text{constant}$, by assumption (a).

(4) $v = \mu\mathrm{E}$ where v is the velocity of migration of a species.

(5) substitute (4) into (3), $\frac{v'_{\mathrm{PTH}}}{\mathrm{E}}-\frac{v'_{\mathrm{DPTU}}}{\mathrm{E}}=\frac{\mathrm{L}}{\mathrm{E}}\left(\frac{1}{t_{\mathrm{PTH}}}-\frac{1}{t_{\mathrm{DPTU}}}\right)=\text{constant}$, by above assumptions where v'_{DPTU} is the velocity of migration of DPTU in the standard mixture of PTHs.

(6) $v=\frac{\mathrm{L}}{t}$

(7) substitute (6) into (5), $\frac{1}{t'_{\mathrm{PTH}}}-\frac{1}{t'_{\mathrm{DPTU}}}=\frac{1}{t_{\mathrm{PTH}}}-\frac{1}{t_{\mathrm{DPTU}}}$ where t'_{PTH} is the normalized migration time and t'_{DPTU} is the migration time of DPTU in the standard.

(8) rearrange (7), $\frac{1}{t'_{\mathrm{PTH}}}=\frac{1}{t'_{\mathrm{DPTU}}}-\frac{1}{t_{\mathrm{DPTU}}}+\frac{1}{t_{\mathrm{PTH}}}$

The columns of time data are corrected as:

(1/norm_time)=(1/std_dptu)-(1/run_dptu)+(1/run_time), where std_dptu and run_dptu are single migration times and norm_time and run_time are columns of migration times. The signal is plotted against norm_time.

CHAPTER 4

DEVELOPMENT OF
A
HIGHLY MINIATURIZED PEPTIDE SEQUENCER

4.1 Introduction

In 1950 Pehr Edman described a method for determining the amino acid sequence of peptides using stepwise degradation with phenylisothiocyanate. Over 40 years later the same chemistry is still used to sequence peptides and proteins but in automated instruments that require less sample *and* time. Even so, researchers are still limited by the minimum quantities of peptide and protein needed to obtain useful sequence information for biomedical research. In this chapter, the design and construction of a highly miniaturized sequencer is described with the goal of sequencing femtomole amounts of peptide and protein. The first section presents the history of automated sequencing and the efforts being made to improve it. The second and third sections describe the construction and characterization of the miniaturized sequencer. Comparison to other instruments is also made in the third section and conclusions follow in the fourth section.

Automated Edman Degradation

Stepwise Edman degradation of peptides and proteins has been a routine tool for primary sequence determination since its introduction in 1950 [1]. The original manual methods of sequencing were soon followed by automated methods, the first of which was described by Edman and Begg in 1967 [2]. Edman and Begg described a liquid-phase sequencer in which peptide or protein was coated as a thin film on the inner walls of a spinning cup. The large surface area of sample was well suited to carrying out coupling and cleavage reactions. Polybrene was later added as a film-stabilizing carrier to help prevent loss of sample during washing steps [3]. In 1971 Laursen [4] described an automated solid-phase sequencer in which peptide was covalently bound to an insoluble resin. The resin was contained in a thermostated glass microbore column and coupling and cleavage reactions were carried out by flowing liquid reagents and solvents through the column. Hewick *et al.* [5] described an automated gas-phase sequencer in 1981, in which peptide or protein was embedded in a matrix of Polybrene. The Polybrene was coated onto a porous glass fibre and sandwiched in a flow-through cartridge, effectively immobilizing the sample. Coupling and cleavage reactions were carried out by flowing gas phase reagents and liquid solvents through the cartridge.

The peptide/protein sequencer described by Hewick *et al.* was capable of giving useful sequence data from as little as 5 picomoles of protein and is considered to be "one of the most salient achievements in the instrumentation of automatic Edman degradation since the report of Edman and Begg in 1967" [6]. While gas-phase sequencing made inroads on the drive to develop micro-sequencers, there is still a need to work toward sequencing samples that are only available in amounts less than one picomole. As Kent *et al.* point out:

> Existing protein sequencing technology allows the determination of amino-terminal sequences from picomole amounts of proteins. Typically, as little as a few tens of picomoles of sequenceable material will allow determination of the sequence of 20 or more residues. However, many proteins of great biological interest can only be isolated in much smaller amounts. This limitation prohibits amino acid sequence determination to characterize the molecules and to assist in further analysis at the DNA level. Protein markers of degenerative nervous diseases in humans are notable examples of such proteins. It has been shown that spinal fluid samples from patients suffering from a number of degenerative nervous diseases contain elevated levels of protein species characteristic of a specific syndrome. This has served as an impetus to develop methods to detect, compare and quantitate nanogram amounts of proteins and has led to the development of such widely used methods as two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and silver staining.
>
> Typically, such rare proteins are present in amounts of 1-10 nanograms of each species in 2D-polyacrylamide gels. For a 33 kDa molecule, this corresponds to 30-300 femtomoles of protein. This is approximately 100- to 1000-fold less than the amount required for protein sequencing using currently available technology, even assuming an ideal isolation method. To study

> proteins at these levels we will have to be able to isolate nanogram amounts of proteins in high yields in a form directly suitable for protein sequence determination and achieve a 1000-fold increase in the sensitivity of sequence determination over methods currently in use [7].

The authors of the quotation above suggest improvements for sequencing ranging from better immobilization of the sample to better data analysis of the products. It is important to improve *all* steps that lead to knowledge of the primary sequence, and develop better methods or steps where possible, as discussed in section 3.1 of Chapter 3. One particular procedure or step that undergoes continuous improvement and optimization is the instrumentation in which phenylisothiocyanate (PITC) degradation and phenylthiohydantoin (PTH) identification are carried out. The desire to construct an instrument that will sequence less than 1 pmol of peptide or protein is shown by the number of recent, related patents [8-19].

The gas-phase sequencer design is still the industry standard for routine peptide and protein sequencing. Commercialized versions of this sequencer are manufactured by Applied Biosystems (ABI), Porton (Beckman), Knauer GmbH, Jaytee Biosciences, Chelsea Instruments and MilliGen [20]. Each of these companies continues to improve their instruments, mostly by miniaturization [21], better chemical compatibility of moving parts [22], and improved detection of degradation products [23]. The large companies are not alone though in the drive to improve the peptide/protein sequencer. For example, a major improvement of the Hewick sequencer was described in 1987 at Caltech, by some of the original researchers [7]. The Caltech sequencer requires less reagents, valves without vacuum-assist, and contains a versatile microprocessor based controller. This sequencer has been used successfully with covalently attached peptides (solid-phase method) and with DABITC double coupling [24]. Other groups have improved the Applied Biosystems Sequencer by modifying sections of the sequencing program, analysing a larger proportion of the PTH amino acid product and designing a smaller reaction cartridge to reduce cycle times [25-28].

Shively's group at the Beckman Research Institute designed a continuous flow reactor (CFR) containing Polybrene-coated silica gel for automated sequencing [29] and multiple manual gas-phase sequencing [30]. The CFR was made of Teflon tubing and fit directly into the commercial automated sequencer, in place of the glass cartridge. Total cycle time therefore depends on the instrument, 60 minutes for the Applied Biosystems instrument or 42 minutes for the Beckman gas-phase instrument [31]. While inexpensive and essentially disposable, Shively suggests that at least 10 pmol sample be applied to the CFR [29]. In the manual gas-phase sequencer, six samples can be sequenced at the same time; however, multiplexing requires 12 valves for flow control. Each additional sample requires two new valves, making the plumbing cumbersome. The total time for coupling and cleavage was only 20 minutes, slightly faster than in other instruments. The anilinothiazolinone (ATZ) amino acid product from the manual sequencer was collected in a 200-μL-size vial, vacuum dried, and manually converted to the PTH form for identification.

A further generation of Shively's instrument combines the continuous flow reactor with a simplified conversion flask, hexagonal valve for fluid inputs, and sonic flow detector for injecting 90% of the PTH product into the HPLC [21]. In this instrument, sample was spotted onto a strip of polyvinylidene difluoride (PVDF) that had been treated with Polybrene. The strip was then inserted into the tubular-shaped continuous flow reactor and reconnected to the hexagonal liquid delivery valve. Each cycle takes 50 minutes, coinciding with the PTH determination time by HPLC. High sensitivity sequence analysis at the 5-10 pmol level was reported [21].

In many cases, peptide or protein is purified by gel electrophoresis; the resulting band containing only a few picomoles of sample. Strategies for electroblotting peptides and proteins directly onto a support that fits into the gas-phase sequencer cartridge have proved quite successful [32-35]. Calaycay *et al.* [21] suggest that their high-sensitivity instrument would be ideal for sequencing peptides and proteins obtained from electrotransfer experiments. Solid-phase sequence analysis has often been used for small peptides to prevent wash-out by solvents and reagents and is well known for improved sensitivity with such peptides. A renewed interest in solid-phase sequencing seems to be the result of improvements in peptide and protein purification by SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis). Electroelution of polypeptides from gels onto covalent coupling resins or membranes allows sequence analysis of picomole amounts of sample [36-38]. Solid-phase methods of peptide immobilization that are compatible with the gas-phase sequencing instrumentation have also been demonstrated [24, 39, 40].

Looking at the history of peptide/protein sequencing, there have been significant impovements in the purification of sample materials and in instrumentation since Hewick's gas-phase sequencer was described in 1981. Each one of these improvements has not led, though, to sub-picomole sequencing on its own. By building on Hewick's original design and detecting products by capillary electophoresis with improved sensitivity, a highly miniaturized sequencer can be developed. This system is described in the following sections.

4.2 Experimental

Instrument Design

A schematic of the highly miniaturized peptide sequencer is shown in Figure 4.1. Reagents and solvents are delivered from the vessels labeled R1-S3 by pressurizing them with argon from either manifold B or C. These reagents/solvents are coupled through valve A to the reaction chamber, I. Coupling and cleavage are performed in the reaction chamber and the ATZ amino acid product is collected in a microcentrifuge vial, K, for conversion. The resulting PTH amino acid is dissolved in 1 µL of electrophoresis running buffer and determined by micellar electrokinetic capillary chromatography (MECC) with thermo-optical absorbance detection. Details of each step are described below.

The reaction chamber, shown in Figure 4.2, is analogous to the reaction cartridge designed by Hewick *et al.* [5]. It is constructed from two pieces of polyimide coated fused silica capillary (Polymicro Technologies) such that a shelf of porous PTFE (Teflon) called Zitex (ABI) supports a glass fibre filter disk (Whatman) onto which the peptide sample is immobilized. The glass fibre is pre-conditioned with Polybrene. A 4-cm length of 400-µm-ID/525-µm-OD capillary and a 12-cm length of 75-µm-ID/360-µm-OD capillary are carefully cut to very flat ends using a fused silica cutting stone (Chromatographic Specialties). The wider capillary is placed perpendicular to the glass fibre filter, gently pushed in like a cork borer, and a tiny mat is pushed into place using the narrower capillary. A small disk of Zitex is inserted in the same manner. The mats are secured in place by epoxying the two capillaries together. Sample introduction is described later in the section called Edman Degradation. Reagents and solvents flow through the reaction chamber during coupling and cleavage.

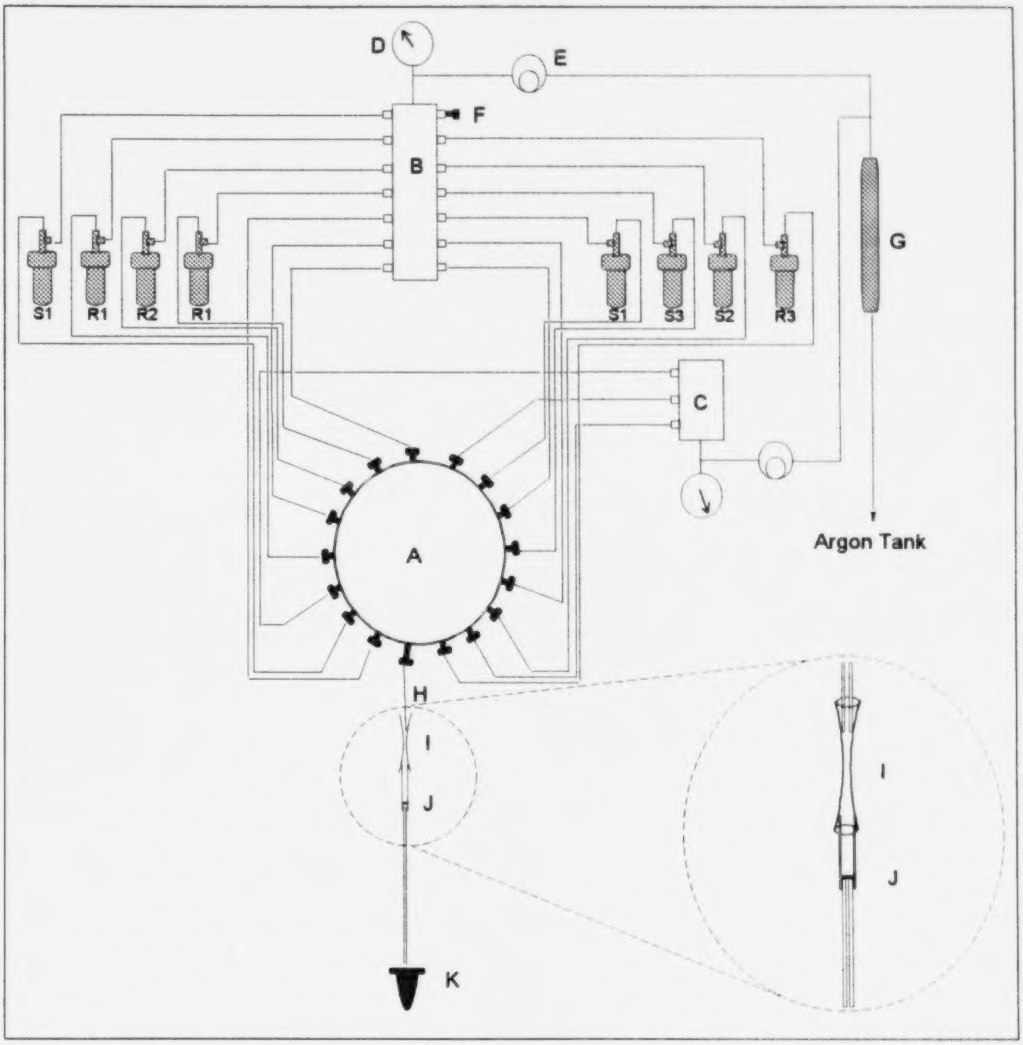


Figure 4.1 Highly miniaturized peptide sequencer schematic. Parts are labeled here and described in detail in the following paragraphs; A: Valco multiposition distribution valve, B: 3.5 psig argon manifold, C: 8 psig argon manifold, D: pressure gauge, E: low pressure regulator, F: pressure relief valve, G: oxygen removal trap, H: outlet capillary, I: fused silica Inner-Lok, J: reaction chamber, K: microcentrifuge vial, R1: 12.5% TMA, R2: 3% PITC in heptane, R3: anhydrous TFA, S1: ethyl acetate, S2: benzene, S3: 1-chlorobutane.

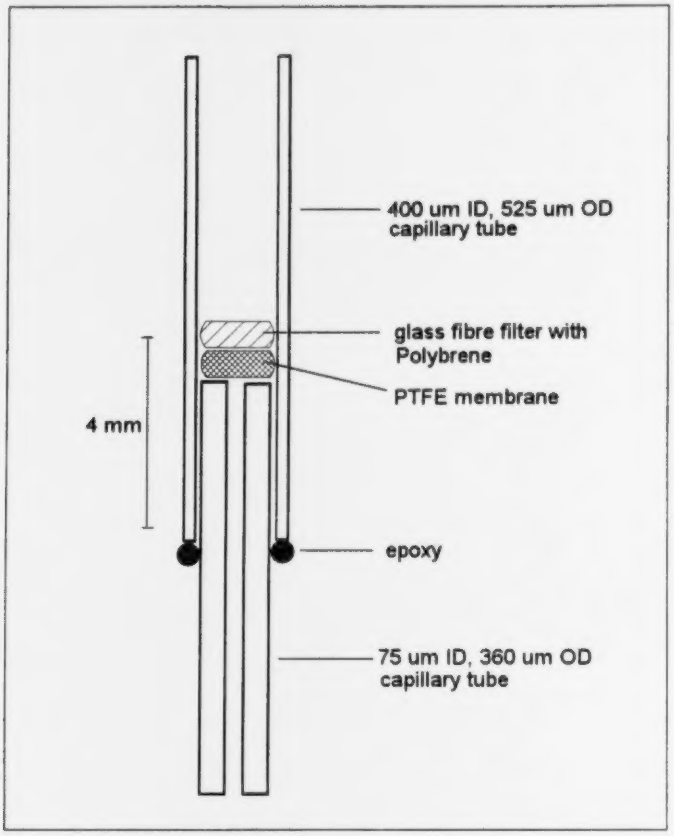

400 um ID, 525 um OD capillary tube
glass fibre filter with Polybrene
PTFE membrane
4 mm
epoxy
75 um ID, 360 um OD capillary tube

Figure 4.2 Diagram of reaction chamber.

The elevated temperature required for coupling and cleavage reactions is provided by solid-state thermoelectrics (Melcor). Thermoelectric devices are made from alternating posts of semiconductor, primarily bismuth telluride, heavily doped to create an excess (N type) of electrons in one post and a deficiency (P type) of electrons in the neighbouring post. The posts are connected in series electrically and in parallel thermally to form a fence-like structure with the top designated as the cold junction and the bottom as the hot junction. Application of a DC voltage creates a carrier current passing through the circuit. Heat absorbed at the cold junction is pumped to the hot junction at a rate proportional to the carrier current and to the number of posts.

Typically thermoelectric devices are used for cooling; however, switching the negative and positive voltage inputs provides a fast and stable method of heating. Control of the temperature is achieved by adjusting the voltage applied to the thermoelectrics with a low voltage, high current power supply (Kepco).

The reaction chamber is dabbed with thermal compound (Dow-Corning) and sandwiched between two copper plates onto which the thermoelectric devices have been glued. Water cooled, copper heat sinks are glued to the opposite side of the thermoelectrics. Heat conducting glue is made by mixing a small amount of silver paint with epoxy. The thermoelectric/copper modules are mounted on a hinged clip, like a clothes peg, so that the reaction chamber can be easily inserted into a machined groove. A small hole (0.5 mm ID) is drilled in one copper plate and a thermocouple (Digi-Sens, Cole Parmer) inserted to monitor temperature during degradation. Figure 4.3 illustrates the thermoelectric/copper module and its orientation with respect to the reaction chamber.

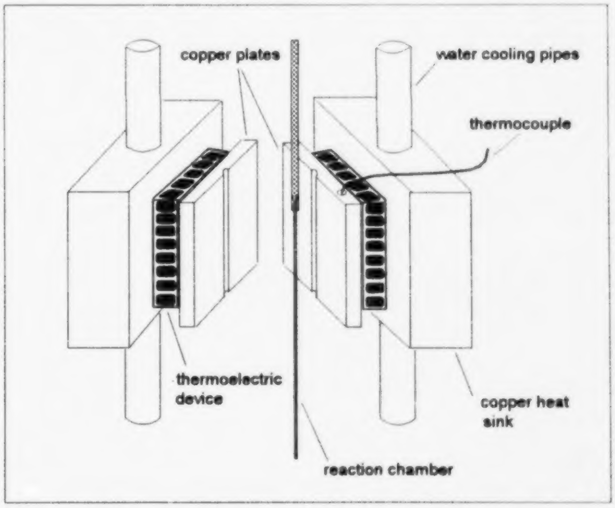


Figure 4.3 Thermoelectric modules for temperature control during coupling and cleavage reactions.

Reagents and solvents required for the coupling and cleavage steps of degradation are delivered from argon (pre-purified, Linde) pressurized vials to a 16-position distribution valve (Valco). Capillary (Polymicro Technologies) or Teflon (Mandel) tubing is used from the vials to the Valco valve. The reaction chamber is connected to a single outlet capillary tube from the valve (see Figure 4.1). Table 4.1 lists the length, size and type of tubing used to direct reagents and solvents to the Valco valve. The outlet capillary (12 cm of 250-μm-ID/340-μm-OD, fused silica) is connected to the wider capillary of the reaction chamber by simply inserting each end into a fused silica compression-type fitting called an Inner-Lok (Polymicro Technologies). As long as the capillary ends are cut flat, an excellent seal is made against the walls of the Inner-Lok because of its tapered construction.

Reagent/solvent delivery vessels are adapted from amber, 5-mL Reacti-vials (Pierce) by machining Teflon inserts (threaded inside with 1/8 inch pipe thread) that fit through the septum holes in the Pierce Reacti-vial caps. Teflon 1/8 inch male-run Ts (Cole Parmer) are screwed into the Teflon inserts; one arm of the T is for argon in, the other arm for reagent out. Figure 4.4 shows the basic configuration of a vessel used for liquid delivery. A Plexiglas rack was machined to hold the Pierce caps in place such that vials hang below, allowing easy removal to refill them.

Argon for pressurizing the reagent/solvent vials is distributed from a stainless steel manifold, machined with 14 outlets, through Teflon lines to the vessels, and maintained at 3.5 psig using a low pressure regulator (Air Products). One of the 14 outlets is fitted with a pressure relief valve (Swagelok). The three outlets pressurizing the TMA and TFA vials are adapted with 0.3 psi check-valves (Swagelok) to prevent back flow of gaseous reagent into the argon manifold. A second manifold at 8 psig is used to deliver argon for some of the drying steps of the degradation procedure.

Manipulation of argon pressure, capillary length, and inner diameter can be used to control flow rates of reagents and solvents to the reaction chamber. Flow control is discussed in detail in Section 4.3 of this chapter. The delivery time for each reagent and solvent is controlled by the multiposition distribution valve.

Table 4.1 Length, size, and type of tubing used to deliver degradation reagents.

Reagent/Solvent	Length to valve	ID (μm)	OD (μm)	Tubing type
Argon	60 cm	75	360	fused silica
12.5% TMA vapour	60 cm	75	360	fused silica
3% PITC in heptane	55 cm	250	340	fused silica
Argon	50 cm	75	360	fused silica
12.5% TMA vapour	60 cm	75	360	fused silica
Argon, 8 pisg	50 cm	250	340	fused silica
Ethyl Acetate	50 cm	250	340	fused silica
Argon	52 cm	75	360	fused silica
TFA vapour	43 cm	250	1588	Teflon
Argon, 8 psig	50 cm	250	340	fused silica
Benzene	47 cm	250	340	fused silica
Argon	50 cm	250	340	fused silica
Cl-butane	50 cm	250	340	fused silica
Argon	58 cm	250	340	fused silica
Ethyl Acetate	60 cm	250	340	fused silica
Argon, 8 psig	50 cm	75	360	fused silica

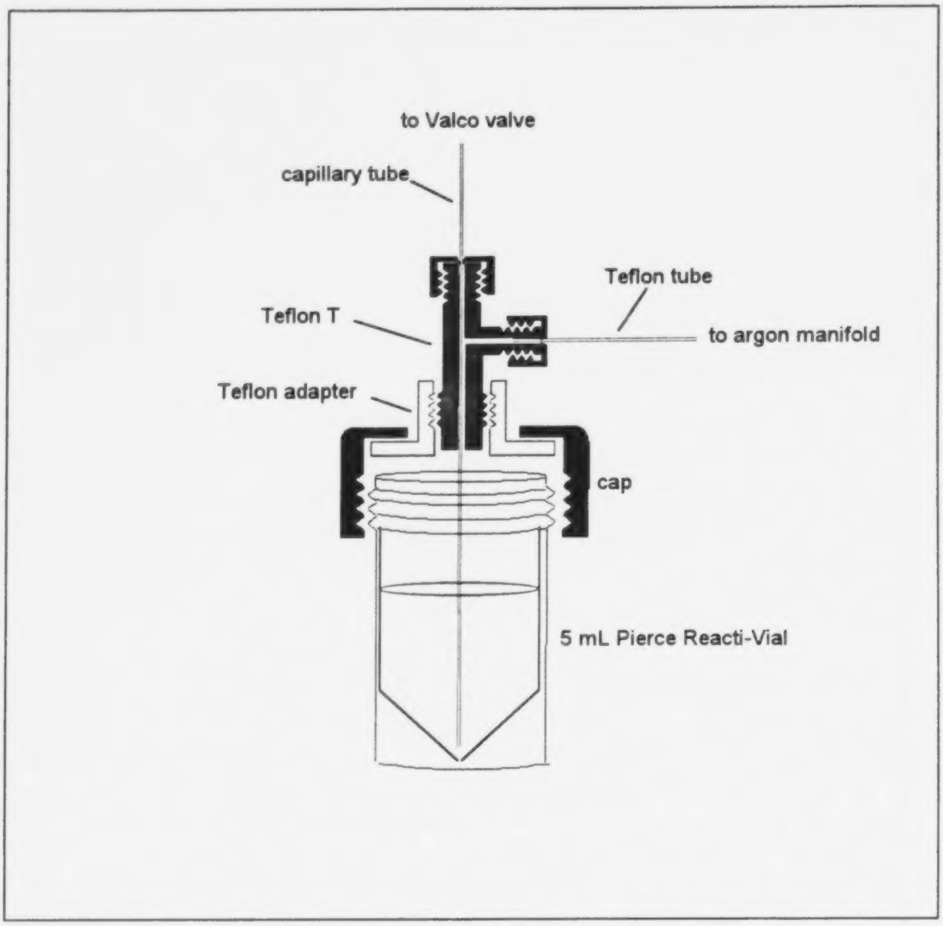


Figure 4.4 Configuration of reagent and solvent delivery vials.

A Valco 16-position distribution valve is used to direct reagent and solvent streams into a single outlet connected to the reaction chamber. Fused silica tubing adapters, 1/16 inch, are used to connect the capillary lines, providing zero dead volume fittings. The valve body is constructed from Hastalloy C and the valve rotor from an inert polymeric material, thus providing reasonable chemical compatibility with Edman degradation reagents. A cross-section of the valve flow path and fitting construction and a diagram of the rotor configuration are shown in Figure 4.5. The valve is mounted on an electric actuator (Valco) that is controlled either by a remote switch or by an RS-232 data interface. The valve is actuated in one direction only, simplifying the sequencer program but limiting the number of possible events to sixteen.

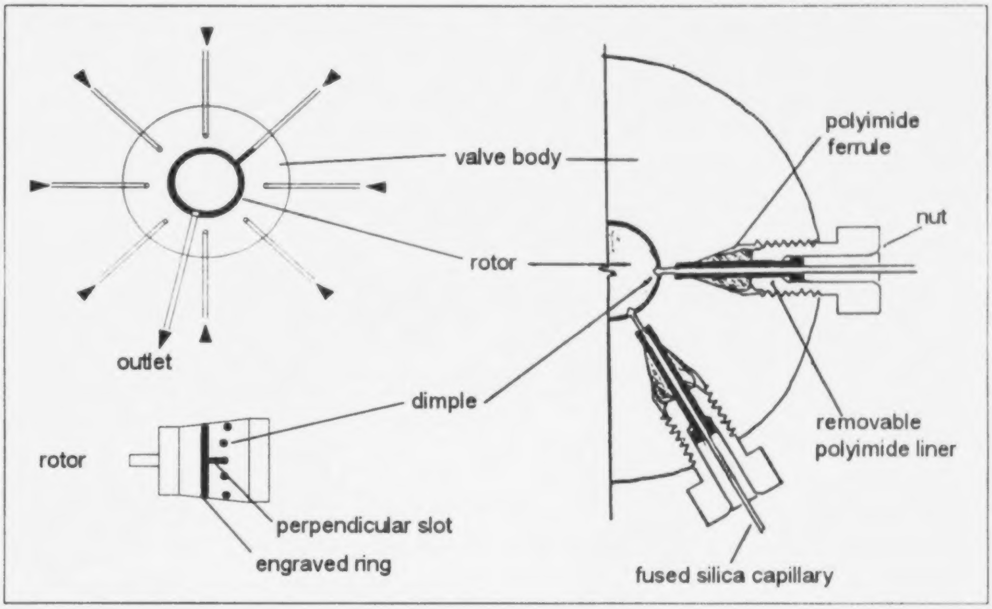


Figure 4.5 Valco multiposition distribution valve.[1]

Edman Degradation

The reaction chamber is easily disconnected from the sequencer by gently pulling it out of the Inner-Lok so that sample can be loaded. A 7 to 10 cm piece of capillary of 75 μm ID/ 360 μm OD is held in a cool flame to burn the polyimide coating off, and then cleaned with optical grade methanol, to make a sample loading capillary. The loading capillary is dipped into the peptide solution for a second and approximately 0.04 μL are drawn up to fill about 1-cm of the capillary in length. The peptide solution contained 2×10^{-3} M insulin chain B in propanol-water, corresponding to 80 pmol in the 0.04 μL amount. The sample loading capillary is inserted into the wide end of the reaction chamber and the peptide solution dispensed onto the glass fibre mat by gravity or by gently pushing air through the sample loading capillary with a syringe. The sample loading capillary is removed, the contaminated end broken off, and the capillary cleaned so it can be reused. The reaction chamber is re-attached to the sequencer by gently pushing the wide end into the Inner-Lok and then sliding the section with the mats into the groove between the thermoelectric modules. The peptide solution is dried by passing argon through the reaction chamber for 5 minutes.

---

[1] Diagrams redrawn from the Valco Instruments Company Inc. technical notes and installation guide.

The reagents and solvents used in the degradation are similar to those used in the 1981 sequencer design by Hewick *et al.* [5]. Coupling is performed by delivering approximately 4 μL of 3% PITC (Sigma, sequencing grade) in heptane (Beckman, sequencing grade) to the reaction chamber, then flowing 12.5% TMA in water (ABI) vapour through at 65°C. Cleavage is performed by flowing anhydrous TFA (Sigma, sequencing grade) vapour through the reaction chamber at 48°C. The ATZ derivatized amino acid product is extracted with approximately 10 μL benzene (Beckman, sequencing grade) and is collected into a 600-μL-size microcentrifuge vial that contains 10 μL 25% TFA. Table 4.2 summarizes the steps for each degradation cycle (excluding conversion), by listing the reagent/solvent, corresponding Valco valve position, length of time the valve is open, and approximate amount or flow rate delivered to the reaction chamber. Sequencing grade ethyl acetate (Beckman) is used twice as a washing solvent and chlorobutane (Fisher) is used to remove residual TFA.

The extracted ATZ amino acid in benzene and aqueous acid are mixed by vortex then reduced to dryness in a vacuum centrifuge (Savant Speed Vac). Manual conversion is performed by adding 20 μL 25% TFA to the dry ATZ amino acid sample, mixing by vortex, and placing the capped microcentrifuge vial in a water bath at 60°C for 20 minutes. The converted PTH amino acid is dried and re-dissolved in 1 μL electrophoresis running buffer for determination. Micellar electrokinetic capillary chromatography (MECC) with laser-based thermo-optical absorbance detection is used to identify the cleaved residue, as described in detail in Chapters 2 and 3.

Table 4.2 Sequencer program.

Valve Position (Step No.)	Reagent or Solvent	Delivery Time (seconds)	Flow Rate or Amount Delivered
1	Argon, 3.5 psi	120	1.6 μL/s
2	12.5% TMA vapours	30	1.6 μL/s
3	3% PITC in heptane	1	3.5 μL
4	Argon, 3.5 psi	90	1.6 μL/s
5	12.5% TMA vapours	420	1.6 μL/s
6	Argon, 8 psi	90	19 μL/s
7	Ethyl Acetate	10	13 μL
8	Argon, 3.5 psi	120	1.6 μL/s
9	anhyd. TFA vapours	300	6.5 μL/s
10	Argon, 8 psi	120	19 μL/s
11	Benzene	15	10 μL
12	Argon, 3.5 psi	120	5.8 μL/s
13	Cl-Butane	60	14 μL
14	Argon, 3.5 psi	150	5.8 μL/s
15	Ethyl Acetate	90	18 μL
16	Argon, 8 psi	120	4.3 μL/s

4.3 Results and Discussion

This section is split into two parts: instrument design and Edman degradation. The first part describes details about the miniature sequencer construction compared to commercially available instruments. The second part describes the results of using the instrument for sequencing a peptide and discusses the observed results on the basis of what other researchers have seen in similar experiments.

Instrument Design

The miniaturized sequencer reaction chamber (Figure 4.2) is analogous to the Hewick design in that it is a flow-through design with the peptide physically immobilized, via Polybrene, on a glass fibre mat. This method contrasts with the spinning cup design or with sequencers designed to chemically immobilize the peptide by covalent attachment (solid-phase design). The difference in reaction chamber volume is dramatic: 150 μL for the Hewick design versus 0.2 μL for the miniaturized design. The diameter sizes of the capillaries used to construct the reaction chamber were chosen based on commercial availability. It takes approximately fifteen minutes to build five reaction chambers, including time for the epoxy to dry. The price of fused silica capillary ranges from $2.50 to $10.00 per meter and, since so little is used for each chamber, the cost is less than $2.00 per reaction chamber. The commercial Zitex membranes are usually about 20 mm in diameter so hundreds of disks of 0.4 mm diameter can be punched out of one piece of Zitex. Similarly, hundreds of glass fibre mats can be made from a single commercial mat. A Polybrene precycled mat of 12 mm diameter, prepared on an Applied Biosystems 477A sequencer, was used in this work and stored at -20°C when not being used to make reaction chambers. The glass fibre mat has poor structural integrity, especially when it is so small, and must be supported by the Zitex mat during sequencing to prevent slippage.

The reaction chamber fits snugly in the machined groove between the copper plates of the thermoelectric modules as shown in Figure 4.3. A small amount of thermal compound ensures excellent thermal contact between the reaction chamber walls and copper plates. Acquiring the desired temperature for coupling or cleavage is achieved by adjusting the voltage applied to the modules. Calibration of the voltage was done by inserting the thermocouple directly into the reaction chamber and measuring temperature versus applied voltage. Figure 4.6 shows the calibration curve.

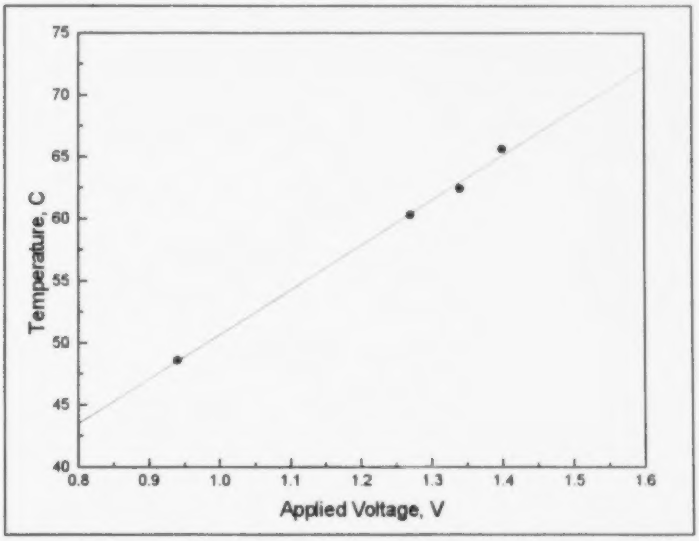


Figure 4.6 Voltage applied to thermoelectric modules versus temperature inside reaction chamber.

Voltage is only applied to the thermoelectric modules during coupling and cleavage steps. The maximum temperature is reached in about 90 s, therefore the voltage is turned on about one minute before the coupling or cleavage begins. For example, the voltage is pre-set at 1.40 V and turned on halfway through Step 1 (see Table 4.2), then turned off after Step 5. All other solvent-wash, argon-dry, and extraction steps are performed at ambient temperature. Thermoelectric devices do not require any servo system or thermostating and remain at the set temperature with fluctuations of only ±2°C. The temperature rise inside the reaction chamber versus time, for a voltage of 0.94 V and maximum current of 2.00 A, is shown in Figure 4.7. The actual current measured, when the power is on, is 0.30 A. Within 30 s of turning the voltage off, the temperature is approximately 28°C.

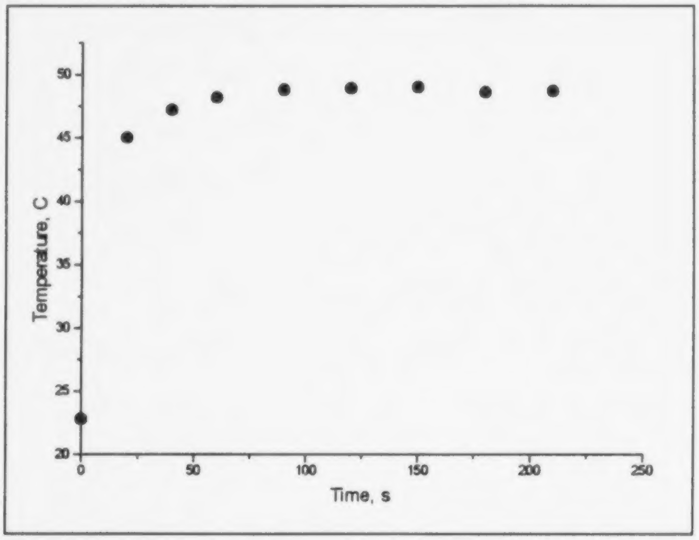


Figure 4.7 Graph of temperature versus time after voltage is applied. The voltage is set at 0.94 volts and the current maximum is set at 2.00 amperes.

The advantage of temperature control with solid-state devices, such as the one used in the miniature sequencer, is that the temperature can be easily changed between the coupling and cleavage reaction steps. In the Applied Biosystems Model 470A instrument, 48°C must be used for coupling *and* cleavage because it is too difficult to change the temperature between steps. A higher coupling temperature would be preferable (55-60°C) but cleavage above 48°C tends to aid acid-catalysed random cleavage of peptide bonds, raising background levels in the PTH residue determinations [41]. In the miniature sequencer design, a higher coupling temperature and lower cleavage temperature can be used due to the rapid heating and cooling times of the thermoelectric devices. Later models of the Applied Biosystems sequencer (Models 477 and 473) have better flexibility of temperature control for coupling and cleavage.

The glass fibre mat in the miniature reaction chamber has an area of 0.4 mm^2. This area is 900 times less than in the Hewick design so it was assumed that 900 times less reagent and solvent could be used in the miniature sequencer. While this reduction may not be appropriate for all the steps in the sequencing program, it provided a starting point from which to optimize the degradation conditions. The reagent/solvent vessels were adapted from 5-mL vials specifically to obtain a large enough cap for a "T" to fit through. Considering the capillary length from vessel to distribution valve and amount used per cycle, a solvent such as ethyl acetate can last for 50 cycles and reagent such as PITC in heptane can last for 990 cycles. Some limitations exist though in the amount of reagent or solvent deliverable.

First, almost all reagents and solvents are pressurized by one argon manifold. A single controlling pressure is used for simplicity of the plumbing. Some commercial sequencing instruments have more than ten different valves and pressure regulators, which all add to the complexity and cost. In the miniaturized sequencer design, once the pressure is set, the tubing dimensions, which vary along the path, determine flow rate. As with the reaction chamber, the diameters of the capillaries from reagent and solvent vessels to the distribution valve were chosen based on standard commercial sizes rather than custom sizes. This choice helps to keep the cost low.

Second, the viscosities of each reagent and solvent vary widely so that at a given argon manifold pressure, flow rates are highly variable. While Poiseuille's Law can normally be used to calculate flow through a tube, the flow path in the miniaturized sequencer is not uniform. Equation 4.1 is a rearrangement of Poiseuille's Law in terms of volumetric flow, F (measured in μL/s), rather than linear velocity:

$$F = \frac{6.9 \times 10^{7} \Delta P \pi r^{4}}{8 \eta L} \tag{4.1}$$

where ΔP is the pressure drop, measured in psi, across a tube of length L (in cm) and radius r (in cm) for a fluid with a viscosity of η (in poise). For ethyl acetate flowing through a capillary with ID=250 µm, L=50 cm, and pressurized at 3.5 psi, the calculated flow rate is 10 µL/s. However, the reagents and solvents do not flow *only* through the capillaries to the distribution valve. Flow must take a tortuous path through the Valco valve (see Figure 4.5), then pass through the glass fibre and Zitex mats, and then through a length of 75 µm ID capillary, which will be the limiting factor according to Equation 4.1. For ethyl acetate flowing through a capillary with ID=75 µm, L=12 cm, and pressurized at 3.5 psi, the calculated flow rate is only 0.35 µL/s, using Equation 4.1.

Third, the rotor configuration in the distribution valve (see Figure 4.5) makes delivery of small volumes (less than 0.5 µL) very difficult. The volume inside the Valco valve was measured as approximately 8 µL and, according to technical information, has channels larger than about 250 µm in diameter. Since there are two possible routes from inlet to outlet, small volumes of liquid tend to get trapped in the valve because the argon, intended to blow out the liquid, takes the easiest route to the valve outlet. Initially, a 12-port air-actuated distribution valve was used to direct reagent/solvent flows. An air-actuated valve was chosen because it can be electrically isolated and moves very fast to the neighbouring position or port. The 12-port valve had an interior (dead) volume of approximately 15 µL and it was difficult to get liquid volumes of less than 2 or 3 µL to blow out with argon. To solve this problem a vacuum line after the reaction chamber was installed; other sequencer designs incorporate vacuum-assisted solvent removal [5].

An additional problem was cross-contamination of reagents and solvents from port to port because the valve rotor was machined with small dimples or divots at each position, as seen in the bottom left of Figure 4.5. These dimples were essentially little wells that filled up with reagent or solvent at each inlet position. Actuation of the rotor would rotate the wells, and liquid in them, to be mixed with reagent in the neighbouring position. This problem was discovered after several months when beads of liquid were observed in an argon delivery capillary after the argon manifold was vented to room pressure. Since it was desirable to have more steps in the sequencing program anyway, a 16-port distribution valve was custom ordered without dimples in the rotor. There was also a problem with the air actuated valve becoming mis-aligned so an electric actuator was attached to the 16-port valve. Electrical actuation is not as fast as air actuation; a half second is about the shortest length of time for the valve to be open to a given port. When the 16-port valve was installed, the vacuum line was removed.

Determining flow rate is made more difficult by the fact that liquid reagents/solvents reach the reaction chamber quickly because initially, travel is through wider-bore capillary. When the narrower section of the reaction chamber is reached, flow is highly constricted and slows down considerably. It is therefore necessary to know (1) the time reagent/solvent takes to reach the reaction chamber and, (2) the flow rate after reaching the reaction chamber. The amount of reagent/solvent delivered can then be estimated. The flow rates listed in Table 4.2 were measured using a flow meter and the delivery times (valve open time) take into consideration the time for reagent or solvent to reach the reaction chamber.

In the original Hewick sequencer design, gas-phase reagents were delivered by bubbling argon through 25% TMA in water and through anhydrous TFA, then directing the vapours to the reaction cartridge [5]. In the Applied Biosystems Model 470A instrument, TFA is delivered by blowing argon over the liquid and TMA is delivered by blowing argon over 12.5% TMA in water [42]. Hawke *et al.* [31] note that they saw no difference in delivering coupling base between the bubble-through versus blow-over method, except a change in the partial pressure of the base. In the miniaturized sequencer it is more convenient to use the blow-over method for delivering vapours of TMA and anhydrous TFA. It is interesting to note that in the Applied Biosystems Model 477A instrument, TFA is delivered to the reaction cartridge as pulses of liquid, not as a gas [43]. Gas-phase delivery of TFA is less harsh and probably better than liquid-phase TFA, which may cause sample wash-out or non-specific cleavage of peptide bonds [44].

It is important that all surfaces in the sequencer be inert and therefore commercial instruments use primarily glass, Teflon, Kel-F and Kal-rez. In the miniature sequencer all reagent and solvent lines, and the reaction chamber, were initially made of polyimide-coated fused silica (glass) capillary. When the TFA vessel was removed the second or third time for re-filling, it was apparent that the acid vapour environment was dissolving the polyimide coating from the capillary. Rather than risk contamination of the TFA, a Teflon tube is used to carry TFA from the vessel to the distribution valve. The only other chemical compatibility problem is the Valco distribution valve. The valve is machined from Hastalloy C, which is slightly better than stainless steel, but is still somewhat poor for TFA vapour. The main problem is adsorption of TFA inside the valve. Even though most of the wetted part inside the valve is the inert rotor material that shouldn't react with TFA, extra wash steps were added to the sequencing program to help remove TFA from the valve as well as from the reaction chamber. These changes are discussed more in the following Edman Degradation section. Ideally, a multiposition Teflon valve should be used; however no manufacturer has been found that distributes a Teflon valve with more than 10 ports that can be remotely actuated. The only alternative is to have a Teflon valve custom made, a costly option.

Edman Degradation

Electropherograms of the standard mixture of 19 PTH amino acids with two by-products (ABI), and the result of one cycle of Edman degradation of insulin chain B peptide performed on the miniaturized sequencer are shown in Figures 4.8A and B. Approximately 50 pmol of insulin chain B in 0.03 µL propanol/water was loaded into the reaction chamber and three cycles of the sequencer program (see Table 4.2) were run. The beginning sequence of insulin chain B is phe-val-asn-gln-his (F-V-N-Q-H-) but only the first cycle gave evidence of PTH amino acid (PTH-phe) cleaved from the peptide. The signal-to-noise ratio in Figure 4.8B is excellent and only the one by-product, DPTU, is observed indicating that the possibility for femtomole level sequencing is very good. Obtaining only one successful cycle of degradation means the coupling, cleavage and conversion reactions are not optimized yet.

The recovered PTH product is dissolved in only 1 µL of electrophoresis buffer and the amount injected for MECC separation is about 1 nL. Therefore, only 0.1% of the sample is injected. The PTH-F peak in Figure 4.8B represents 45 fmol, which corresponds to the yield of 90% for the first degradation cycle. Lack of sequence information on subsequent cycles may be caused by various problems, some of which are discussed on the next few pages.

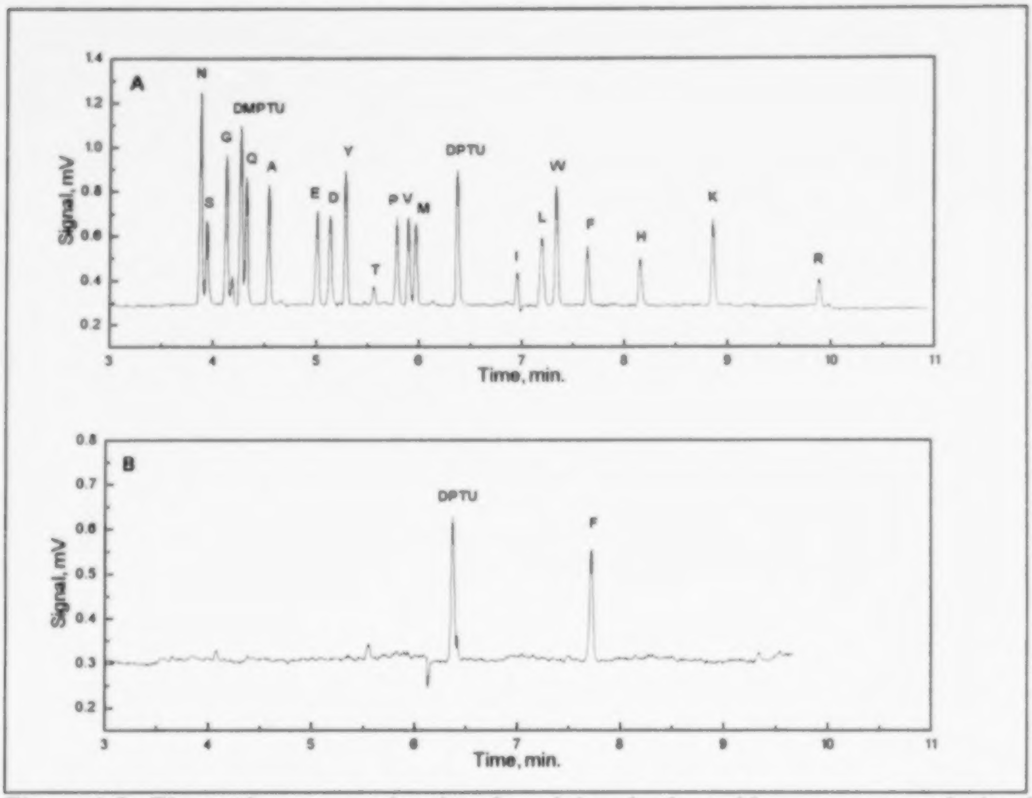


Figure 4.8 Electropherograms showing the miniaturized peptide sequencer analysis of 50 pmol insulin chain B, (A) standard, (B) first degradation cycle. MECC was performed in a 50-μm-ID, 40-cm-long fused silica capillary with detection 35 cm from injection. Electrokinetic injection from 1 μL of product was made at 2 kV for 5 s and separation proceeded at 10 kV in a buffer of 10.7 mM sodium phosphate, 1.8 mM sodium tetraborate and 25 mM SDS, pH 6.7.

Sample loading into the miniaturized sequencer reaction chamber is simple and fast; however only small volumes can be applied. The commercial gas-phase sequencers use a glass fibre filter disk 12 mm in diameter onto which 30 μL of peptide or protein solution can be spotted, usually in 10-15 μL batches. The glass fibre disk in the miniaturized reaction chamber is only 0.4 mm in diameter so proportionately 0.03 μL can be loaded. As described in the experimental section, a capillary tube is used to load this small volume. The peptide solution was sometimes loaded under a microscope to watch delivery. On occasion, part of the liquid would pass *through* the mat suggesting overloading. In these instances, less than 80 pmol is retained. Shively *et al.* [29] note than if the disk is overloaded or overwetted then the sample may be wholly or partly lost. In the failed attempts at obtaining sequence information with the miniature sequencer, perhaps sample was lost because of overwetting.

Several modifications to the sequencer program were made before the steps listed in Table 4.2 were chosen. The choice of reagents and solvents for the miniaturized sequencer were guided by the Hewick gas-phase sequencer protocol. The number of steps in the sequencing program is limited to 16, and was initially limited to 12—the number of positions on the distribution valve. These limitations mean that several short-cuts are taken: only one delivery of PITC instead of two, only one wash after coupling with ethyl acetate and no additional heptane or benzene washes, only one extraction after cleavage instead of two. Hawke *et al.* [31] used a similar sequencing program with only 17 steps and a single PITC in heptane delivery. They did, though, deliver TFA three times separated by drying steps. No solvent washes after TFA delivery were indicated.

Benzene is used to extract the ATZ amino acids after cleavage, rather than 1-chlorobutane, because the future intent is to freeze and immobilize the extracted product for in-line conversion of ATZ to PTH. The PTH product would then be injected directly into a separation capillary for determination. Benzene freezes at 5.5°C whereas chlorobutane freezes at -123°C so benzene was chosen for convenience of freezing temperature. The carcinogenic danger of handling benzene should be minimal since such small volumes are required for sequencing. While the original Hewick sequencer and the early model ABI sequencer used vacuum assisted solvent removal, the newer ABI and Beckman instruments do not, so the vacuum steps were eliminated.

Hunkapiller [41] states that in the Applied Biosystems Model 470A sequencer, TFA delivery is followed by argon drying, which removes most residual TFA. In the miniaturized sequencer, it is very difficult to remove the TFA after cleavage. After more than 5 minutes of argon drying there is still acidic vapour coming out the bottom of the reaction cartridge. A moistened piece of litmus paper held at the chamber exit is very effective for indicating neutral, acidic or basic vapours, or liquid, coming out. The inability to remove TFA efficiently might be a factor in poor sequencing results. Residual TFA washed into the reaction chamber with ATZ extraction solvent (benzene) may cause dissolution of the peptide during the extraction process. Extraction of the ATZ amino acid *and* truncated peptide would account for lack of sequence in the second cycle. In a related problem, when the 12-port valve was being tested, only 10 to 20μL benzene passed through the reaction chamber between cleavage with TFA and delivery of 12.5% TMA in the next cycle. The result was white "smoke" coming out the exit of the reaction chamber, indicating formation of a precipitate (salt) from reaction of residual TFA with TMA. The reaction was obviously undesirable. Additional washing with other solvents to prevent salt formation became possible when the 16-port distribution valve was installed.

The 16-port valve allows more flexibility with sequencing steps, compared to the 12-port valve. After benzene extraction of ATZ amino acid, a chlorobutane wash step was added but by the third cycle, evidence of an acidic environment was seen during the coupling step. An additional ethyl acetate wash step, inserted immediately after chlorobutane, improves removal of TFA. Finally, an argon drying step between each solvent wash gives the best removal of TFA. It is possible that residual TFA in the reaction chamber hydrolyses the truncated peptide after the first cycle. This process could be aided by the high temperature used for coupling and by the high water vapour content of TMA coupling buffer. It may be causing lack of sequence information after the first degradation cycle. Coupling of the peptide with PITC was performed at 65°C. Tarr [45] suggests that above 60°C, excessive by-products and/or blocking reactions will result. It is possible that after one cycle, the amino terminal end of the truncated peptide is blocked, preventing further degradation. Hydrolysis of PITC coupling reagent into aniline is also possible at 65°C. Tarr suggests that a coupling reaction is 99% complete at one minute, which means that the coupling time should be reduced in the miniaturized sequencer. A reduced reaction time would minimize hydrolysis and unwanted side reactions.

The reactive nature of all the Edman degradation reagents is quite apparent by the problems encountered. When basic vapours were observed coming out of argon lines attached directly to the low pressure argon manifold, it was discovered that the Swagelok check valves had failed. Check valves have Viton popets that seal to prevent backward flow of gas. Chemical compatibility tables show extreme reaction of triethylamine (TEA) with Viton so it is presumed that TMA caused deterioration of the popets, rendering the check valves useless. New polyethylene popets are now installed and seem to work well. Hawke [31] *et al.* recommend that conventional O-rings or gaskets should not be used unless they are truly inert. Hawke and co-workers suggest that the best material for these parts is Kal-Rez® from Du Pont.

Delivery of PITC in heptane is problematic. Very small volumes (< 1 μL) cannot be delivered because of the mechanics of the distribution valve. Currently, the miniaturized sequencer delivers 3.5 μL of 3% PITC in heptane, which is equivalent to

830 nmol. This amount is two or three orders of magnitude more than required for coupling to 50 pmol peptide. However, much of the 3.5 μL passes through the mat rather than barely wetting it. Ideally, the heptane should evaporate after delivery, leaving PITC to react with the peptide. It is possible that for some degradation cycles, insufficient PITC is staying on the glass fibre mat. Coupling is incomplete. Reproducibility of reagent delivered must be improved to solve this problem.

It is important to compare the amount of reagents and solvents used in the miniaturized sequencer to other sequencers. Table 4.3 presents a comparison of the miniaturized sequencer to eight other research groups' instruments. In each case except Totty *et al.*, a gas-phase sequencer with a flow-through sample cartridge was used, in both commercial and non-commercial instruments.

The values in Table 4.3 show similar amounts of PITC but greater amounts of TMA and TFA delivered for other instruments compared to the miniaturized sequencer. Too little TMA present during coupling may be causing insufficient buffering of the reaction of PITC with the peptide. Although the first degradation cycle is unaffected, a high concentration of hydrogen ions may be produced during the coupling reaction [47] and if buffer capacity is poor, it leads to hydrolysis of the truncated peptide. The result would be lack of sequence information in cycles two, three, four, and so on. If less PITC could be delivered to the reaction chamber in the miniaturized sequencer, and the flow rate of 12.5% TMA is increased slightly, it should be possible to reduce coupling time. The cycle times range from 42 to 55 minutes per cycle for the instruments shown in Table 4.3. The miniaturized sequencer program described in Table 4.2 requires 31 minutes for coupling and cleavage, similar to other instruments. It is expected that the oupling and cleavage times could both be reduced, especially when samples of less than one picomole are degraded.

Table 4.3 Comparison of reagent/solvent amounts for miniaturized sequencer versus other gas-phase sequencers.

Reagent or Solvent	This Work	Hewick *et al*. (1981) [5]	Esch (1984) [46]	Hawke *et al*. (1985) [31]	Haniu *et al*. (1988) [30]	Hunkapiller (1988) [41]	Baumann (1990) [26]	Calaycay *et al*. (1991) [21]	Totty *et al*. (1992) [27]
μL PITC [a]	0.1	7.5	75	1.0	0.25-0.5	3.6	--	0.1	1.5
TMA[b] flow time amt.	1.6 μL/s 420 s 0.67 mL	83 μL/s 900 s 74.4 mL	-- 1200 s --	-- 600 s 0.040 mL [f]	33 μL/s 300-600 s 9.9-20 mL	40 μL/s [d] 1200 s 48 mL	-- 1200 s --	-- 690 s --	-- 800 s --
TFA flow time amt. with Ar	6.5 μL/s 300 s 1.95 mL	83 μL/s 650 s 53.5 mL	-- 1670 s --	-- 810 s 0.25 mL [f]	33 μL/s 180 s 5.9 mL	233 μL/s [e] 760 s 177 mL	140 μL/s 650 s 91 mL	40 μL/s [d] 500 s 20 mL	liquid delivery
ATZ extraction solvent [c]	10 μL	1200 μL	1200 μL	1000 μL	200 μL	130 μL [d]	--	240 μL	150 μL
25% TFA	20 μL	50 μL	50 μL	20 μL	20 μL	48 μL [d]	100 μL	60 μL	20 μL [d]

[a] amount of PITC adjusted to reflect % in heptane.

[b] 12.5% or 25% TMA in water.

[c] benzene or chlorobutane.

[d] value estimated from pressurization used for reagent/solvent.

[e] value obtained from reference [26].

[f] amount of liquid consumed per cycle, although reagent delivered in gas phase

-- value not available.

To compare the miniaturized sequencer performance directly to a commercial sequencer, insulin chain B was degraded on the ABI Model 470A gas-phase protein sequencer. Chromatograms of the PTH standard and first, sixth and tenth cycles of degradation of 200 pmol insulin chain B are shown in Figures 4.9A to D. The baseline disturbances that are seen before 5 minutes are caused by the solvent gradient used in the HPLC separation. The main differences that can be seen between MECC and HPLC determination of the degradation products are the total time for separation and the efficiency of separation. The MECC technique takes almost a third less time (10 minutes versus 27 minutes) to separate the standard and total determination time is about 5 times faster; the HPLC system has to flush the column with 90% acetonitrile and then re-establish the solvent gradient after each run. Peak efficiencies are approximately 10,000 theoretical plates for the HPLC separation and at least 250,000 for MECC separations. Peak efficiency and resolution are extremely important here because identification of the PTH residue from degradation is based solely on retention time.

The conversion of ATZ to PTH amino acid takes place in the ABI 470A sequencer in a flask of relatively large volume—1,000 μL. The converted PTH residue is dried then reconstituted in 120 μL of 20% acetonitrile. The sample loop for HPLC injection is 50 μL which means that only 40% of the sample is identified. For the sequencing results shown in Figure 4.9B–D, the PTH-phenylalanine (F) peak from cycle 1 represents 43 pmol, the PTH-leucine (L) peak from cycle 6 represents 21 pmol and the PTH-histidine (H) peak from cycle 10 represents 12 pmol. Considering that 40% of the extracted product is analysed and the initial amount of peptide was 200 pmol, the yields for cycles 1, 6 and 10 are 54%, 26% and 15% respectively. These cycle yields give a repetitive yield of 87% using the calculation shown in Equation 3.1.

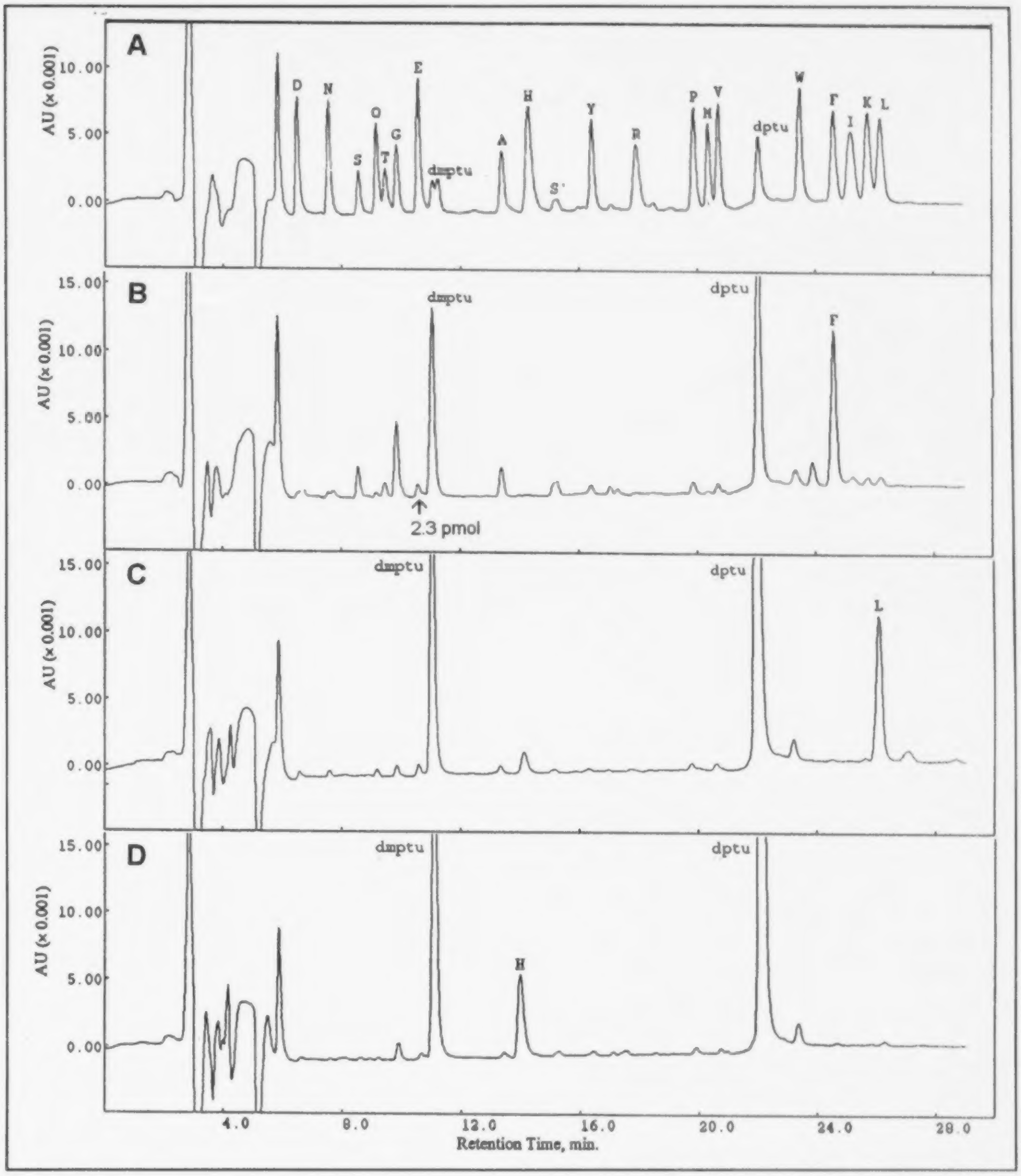


Figure 4.9 Chromatograms showing the ABI Model 470A gas-phase sequencer analysis[2] of 200 pmol insulin chain B, (A) standard, (B) cycle 1, (C) cycle 6, (D) cycle 10. Reversed phase-HPLC separation was performed using a C-18 reversed phase column (Brownlee), 220 mm long × 2.1 mm ID, 5-μm particle size. Solvent A is 50-100 mM sodium acetate buffer with 3.5% THF, pH 4.00 and solvent B is acetonitrile. The separation gradient proceeds from 9-11% B to 35-40% B with a flow rate of 200 μL/min.

---

[2] Sequencing details and chromatograms gratefully obtained from Dr. L.B. Smillie and M.R. Carpenter, Department of Biochemistry, University of Alberta.

The most important difference between the PTH determinations shown in Figures 4.8 and 4.9 is the detection capability. The small peak just before DMPTU in Figure 4.9B is labeled as 2.3 pmol, which is the computer integrated amount based on the standard. Considering that the ABI 470A sequencer only injects 40% of the degradation product, this labeled peak actually represents 5.7 pmol. If individual cycle yields were even 80%, 5 pmol is reasonably the smallest amount of purified peptide or protein that could be sequenced in the ABI 470A instrument because 0.5-1 pmol is about the limit of detection for HPLC identification. MECC with thermo-optical absorbance detection offers almost 100-fold better mass sensitivity for PTH amino acid determination compared to HPLC.

Both the ABI 470A instrument and the miniaturized sequencer suffer from the same limitation—inefficient injection of the degradation products. In the commercial instrument, injecting all of the product will provide a two-fold increase in sequencing sensitivity. Some authors have addressed this problem by using either a flow restriction and valve after the HPLC sample loop to avoid wasting precious PTH product [28] or by using a sensor to indicate when the loop is full [21]. In the miniaturized peptide sequencer, injecting all of the product will provide a thousand-fold increase in sequencing sensitivity. An integrated conversion chamber similar in construction to the reaction chamber could be built with on-line CE separation of the PTH product directly from the conversion chamber. This improvement in identification of the Edman degradation products will help reach the requirements of nanogram-level sequencing set out in the quote by Kent *et al.* [7] at the beginning of this chapter.

4.4 Conclusions

The desire to obtain primary sequence information on smaller and smaller amounts of peptide and protein is apparent by the continued efforts to develop instruments and techniques with this capability. Better sample purification and isolation, better sample immobilization, better instrumentation and better detection of sequencing products are all being pursued to meet the same goal: sequencing less than one picomole. This chapter presented work on improving instrumentation as a means to the sequencing goal: a miniaturized peptide sequencer with capillary electrophoretic determination of the derivatized amino acids.

Problems encountered during development included mechanical failure of the distribution valve, reagent compatibility with construction materials, and reagent/solvent delivery timing. Although TFA was a problem *inside* the distribution valve, an advantage of consuming small amounts of reagents is the reduced chance of corrosion of *other* parts in the sequencer. Several authors [7, 31] emphasize that the chemistry components be well separated or isolated from controlling electronics to avoid corrosion.

Dimethylamine, an impurity in the TMA coupling base, is usually responsible for DMPTU formation during coupling. Very little of this by-product was seen. A lack of large amounts of DMPTU suggests that either too little TMA was used in the reaction or solvent wash steps were very effective. It is possible to use a different coupling base such as 2% TEA (triethylamine) recommended by Calaycay *et al.* [21] or NEM (n-ethyl morpholine), which has a lower pKa and is used in the MilliGen solid-phase sequencer [43] or DIEA (diisopropylethylamine) which is used in the Porton instrument [42]. Either of these two alternative bases would be preferable to the highly odorous TMA in the opinion of the person operating the sequencer or sharing the laboratory space! Quality of the solvents must be explored as well in an attempt to improve the performance of the miniaturized sequencer. Trace impurities of aldehydes or peroxides can cause blocking of the amino terminus, ending sequencing abruptly.

The lack of sequence information in the second and third and further degradation cycles may have been caused by peptide washing out of the reaction chamber. This problem could be checked by using a radiolabeled peptide and recording counts before sequencing and then after a few cycles to see how much, if any, peptide is lost. Improved methods of immobilization must be explored such as using different membranes [35] or using solid-phase methods that covalently attach peptide to a support [38, 39, 48-51]. Methods that have been developed for covalent attachment of peptide to a support after separation by slab gel electrophoresis [24, 36] are important since it will be necessary to sequence electroblotted peptides and proteins in the miniaturized sequencer. Also, covalent attachment of peptides may be necessary because of difficulties in drying the peptide by flowing argon through the capillary.

A few refinements in the sequencing procedure and mechanics are necessary to produce extended sequence information from the miniaturized peptide sequencer. Once these parameters are optimized, in-line conversion of extracted ATZ to PTH amino acid will be implemented. The entire amount of PTH amino acid can then be injected directly from a capillary-sized conversion chamber into the capillary electrophoresis analysis system. MECC with laser-based thermo-optical detection for determination of PTH amino acids provides fast, efficient separation of femtomole quantities of PTH derivatives as demonstrated in Chapters 2 and 3.

4.5 References

1. P. Edman, *Acta Chem. Scand.* **4**, 283-293 (1950).
2. P. Edman and G. Begg, *Eur. J. Biochem.* **1**, 80-91 (1967).
3. G.E. Tarr, J.F. Beecher, M. Bell and D.J. McKean, *Anal. Biochem.* **84**, 622-627 (1978).
4. R.A. Laursen, *Eur. J. Biochem.* **20**, 89-102 (1971).

5. R.M. Hewick, M.W. Hunkapiller, L.E. Hood and W.J. Dreyer, *J. Biol. Chem.* **256**, 7990-7997 (1981).
6. K.-K. Han, D. Belaiche, O. Moreau and G. Briand, *Int. J. Biochem.* **17**, 429-445 (1985).
7. S. Kent, L. Hood, R. Aebersold, D. Teplow, L. Smith, V. Farnsworth, P. Cartier, W. Hines, P. Hughes and C. Dodd, *BioTechniques* **5**, 314-321 (1987).
8. G.F. Begg, R.J. Simpson and A.W. Burgess, Patent WO 9115758, Switz., *Methods and apparatus allowing sequential chemical reactions* (1991).
9. M. Rusnak, J.E. Shively and J.R. Calaycay, Patent WO 9009595, USA, *Apparatus and method for protein and peptide sequencing* (1990).
10. M.W. Hunkapiller, Patent EP 295966, USA, *Method and apparatus for determination of peptide sequences* (1988).
11. K.P. Hupe and F. Strohmeier, Patent EP 279882, Fed. Rep. Ger., *Apparatus for the stepwise preformance of chemical reactions, particularly automatic sequencing of proteins and synthesis of oligonucleotides and/or peptides.* (1988).
12. N. Takagi, Patent JP 62184355, Japan, *Method for primary structure determination of proteins* (1987).
13. H.A. Penhasi and M. Qadeer, Patent WO 8702139, USA, *Method and apparatus for sequening small samples of peptides and proteins* (1987).
14. L.E. Hood, M.W. Hunkapiller, R.M. Dreyer, A.W. Stark and R.M. Hewick, Patent FR 2490505, USA, *Apparatus and method for carrying out chemical processes in sequence* (1982).
15. J.E. Shively, Patent EP 299741, USA, *Protein or peptide sequencing method and apparatus* (1989).
16. A. Tsugita, M. Kamo, T. Uchida, I. Nanno, Y. Nomoto and S. Takahashi, Patent EP 372949, Japan, *Reaction vessel for amino acid sequence analysis* (1990).
17. T. Uchida, Patent US 5135720, Japan, *Reaction chamber for amino-terminal sequence analysis of peptides or proteins* (1992).
18. F. Sakiyama, S. Tsunasawa and T. Harada, Patent JP 62167480, Japan, *An amino acid sequence automatic analyser* (1987).
19. T. Uchida, A. Tsugita, K. Takamoto and K. Satake, Patent EP 529604, Japan, *Degradative method for determination of a sequence of a protein or a peptide* (1993).
20. M.J. Geisow and A. Aitken in *Protein Sequencing, a Practical Approach,* (eds. J.B.C. Findlay & M.J. Geisow) 85-98 (IRL Press; Oxford, 1989).
21. J. Calaycay, M. Rusnak and J.E. Shively, *Anal. Biochem.* **192**, 23-31 (1991).

22. W. Carson, *Personal Communication* (1993).

23. M. Albin, *Personal Communication* (1991).

24. R. Aebersold, G.D. Pipes, R.E. Wettenhall, H. Nika and L.E. Hood, *Anal. Biochem.* **187**, 56-65 (1990).

25. P. Tempst and L. Riviere, *Anal. Biochem.* **183**, 290-300 (1989).

26. M. Baumann, *Anal. Biochem.* **190**, 198-208 (1990).

27. N.F. Totty, M.D. Waterfield and J.J. Hsuan, *Protein Science* **1**, 1215-1224 (1992).

28. G.S. Begg and R.J. Simpson in *Methods in Protein Sequence Analysis, Proceedings of the 7th International Conference, Berlin, July 3-8, 1988,* (eds. B. Wittmann-Liebold) 108-111 (Springer-Verlag; Berlin, 1989).

29. J.E. Shively, P. Miller and M. Ronk, *Anal. Biochem.* **163**, 517-529 (1987).

30. M. Haniu and J.E. Shively, *Anal. Biochem.* **173**, 296-306 (1988).

31. D.H. Hawke, D.C. Harris and J.E. Shively, *Anal. Biochem.* **147**, 315-330 (1985).

32. T. Bergman and H. Jornvall, *Electrophoresis* **11**, 569-572 (1990).

33. T. Choli and B. Wittmann-Liebold, *Electrophoresis* **11**, 562-568 (1990).

34. L.D. Ward, G.E. Reid, R.L. Moritz and R.J. Simpson, *J. Chromatogr.* **519**, 199-216 (1990).

35. D.F. Reim and D.W. Speicher, *Anal. Biochem.* **207**, (1992).

36. R.H. Aebersold, G.D. Pipes, H. Nika, L.E. Hood and S.B.H. Kent, *Biochem.* **27**, 6860-6867 (1988).

37. S.-P. Liang, T.T. Lee and R.A. Laursen, *Anal. Biochem.* **197**, 163-167 (1991).

38. D.J.C. Pappin, J.M. Coull and H. Koester in *Current Research in Protein Chemistry: Techniques, Structure and Function,* (eds. J.J. Villafranca) 191-202 (Academic Press Inc.; San Diego, CA, 1990).

39. V.K. Sarin, Y. Kim and J.L. Fox, *Anal. Biochem.* **154**, 542-551 (1986).

40. J.M. Coull, D.J. Pappin, J. Mark, R. Aebersold and H. Koster, *Anal. Biochem.* **194**, 110-120 (1991).

41. M.W. Hunkapiller in *Protein/Peptide Sequence Analysis: Current Methodologies,* (eds. A.S. Bhown) 87-117 (CRC Press; Boca Raton, 1988).

42. M.R. Carpenter, *Personal Communication* (1992).

43. T. Bures, *Personal Communication* (1992).

44. R. Aebersold, *Personal Communication* (1992).

45. G.E. Tarr, *Methods Enzymol.* **47**, 335-57 (1977).

46. F.S. Esch, *Anal. Biochem.* **136**, 39-47 (1984).

47. P. Edman and A. Henschen in *Protein Sequence Determination*, (eds. S.B. Needleman) 232-279 (Springer-Verlag; New York, 1975).

48. R.A. Laursen, J.D. Dixon, S.-P. Liang, D.M. Nguyen, T. Kelcourse, L. Udell and D. Pappin in *Methods in Protein Sequence Analysis, Proceedings of the 7th International Conference, Berlin, July 3-8, 1988,* (eds. B. Wittmann-Liebold) 61-68 (Springer-Verlag; Berlin, 1989).

49. R.A. Laursen, T.T. Lee, J.D. Dixon and S.-P. Liang in *Methods in Protein Sequence Analysis*, (eds. H. Jornvall, J.-O. Hoog & A.-M. Gustavsson) 47-54 (Birkhaauser Verlag; Basel, 1991).

50. G.E. Tarr, *J. Prot. Chem.* **7**, 293-295 (1988).

51. J.E. Walker, I.M. Fearnley and R.A. Blows, *Biochem. J.* **237**, 73-84 (1986).

CHAPTER 5

CAPILLARY ELECTROPHORESIS AND THERMO-OPTICAL ABSORBANCE DETECTION OF BASES, NUCLEOSIDES, AND DNA

5.1 Introduction

Analysis of complex mixtures of biologically important samples is a fundamental step in medical and biochemical research. In this chapter, the separation and detection techniques described in the first three chapters are applied to the analysis of bases, nucleosides and DNA fragments. The first section provides a description of the analysis method and samples studied as well as an account of the relevant literature. Sections 5.2 to 5.4 present the experimental details, and results of the analyses, and a comparison to other methods.

Capillary Electrophoresis

Capillary electrophoresis (CE) is a separation technique rapidly gaining importance for the analysis of complex mixtures. Besides superb resolution of up to a million theoretical plates, and fast separation in as little as a few seconds, the ability to handle small sample sizes make CE an attractive method for analysing biological samples that may be difficult to isolate and purify in large quantities. Intracellular components, rare peptides and proteins, and forensic samples are ideal for CE separation technology.

Analysis of a nanolitre of material requires innovative detection methods, some of which were mentioned in Chapter 1. The most sensitive method of detection for biomolecules separated by CE is fluorescence. Sub-attomole detection of fluorescently labeled amino acids [1], single-stranded DNA [2], double-stranded DNA [3], monosaccharides [4] and oligosaccharides [5] have been reported. It is not always convenient, though, to label the molecules of interest; the concentration may be too low for the labeling reaction to go to completion or cell lysates may need to be directly analysed with no room for a labeling step [6, 7]. When labeling is not feasible, detection must rely on the spectroscopic properties of the native molecule. Usually, this means absorbance based methods of detection, since many molecules do not fluoresce naturally or have low quantum efficiencies if they do fluoresce [8].

Absorbance in the near UV is the most common method of detection used in CE analysis. However, since absorbance methods are path length dependent, and efficient CE separations require capillaries with $ID < 50\ \mu m$, the detection limits are poor. Limits of detection for CE are typically near the low picomole level for absorbance-based methods. The laser-based thermo-optical absorbance detector, described in Chapters 1 and 2, provides improved detection of derivatized amino acids separated by CE [9-15]. Amino acids are biologically important because they are the constituents of proteins. Another class of biologically important molecules are ribonucleic and deoxyribonucleic acids and their constituents: purine and pyrimidine bases, nucleosides and nucleotides.

Nucleic Acids and Their Constituents

Nucleic acids are the most important components of living cells because they are the chemical carriers of genetic information. Nucleic acids alone represent the potential for self-duplication. In addition, the nucleic acids are responsible for making the the next most important components of living cells: proteins. Deoxyribonucleic acid (DNA) is a large polymeric molecule consisting of two anti-parallel nucleotide strands hydrogen bonded together through the nucleotide bases. Figure 5.1 shows a schematic of the primary structure of DNA where a phosphate-sugar-base group represents one nucleotide, a sugar-base group represents one nucleoside and a base is a heterocyclic amine. Ribonucleic acid (RNA) is similar in structure to DNA and is synthesized (transcribed) from DNA. RNA in turn decodes the genetic message and uses the information to build proteins, a process called translation.

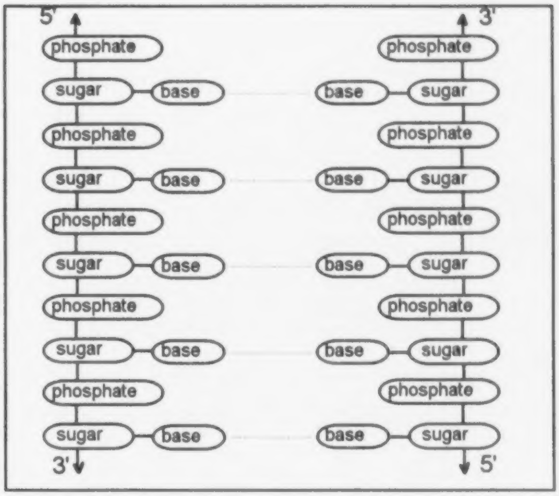


Figure 5.1 Schematic diagram of the primary structure of DNA. The 3' and 5' labels refer to positions on the sugar ring for attachment of phosphate.

In DNA and RNA, there are two purine bases, adenine and guanine, and three pyrimidine bases, cytosine, thymine and uracil. When the bases are attached to ribose rings (sugar), the resulting structures are nucleosides: adenosine, guanosine, cytidine, thymidine and uridine. Attachment of phosphate (H_3PO_4) to the 5' position on the ribose ring completes the nucleotide structures. The structures of DNA and RNA differ slightly in three respects. First, at the 2' position of the ribose ring, DNA has an H rather than an OH group, hence the name *deoxy*ribonucleic acid. Second, DNA does not contain the base uracil and RNA does not contain the base thymine. Third, RNA exists only as a single strand whereas DNA exists as single- or double-stranded (Figure 5.1).

Besides the inherent importance of bases, nucleosides and nucleotides as building blocks for DNA and RNA, analogs of nucleic acid derivatives have interest as therapeutic agents, as tags, or as modifiers in biochemical studies. For example one of the most common drugs used in AIDS treatment is AZT: azidothymidine [7], an analog of thymidine. Separation and identification of these components is important for determining abnormalities that may cause illness or for determining up-take of a drug used to combat illness. Workers in Zare's group [7] analysed intracellular free ribonucleotides, using capillary zone electrophoresis in order to obtain an intracellular nucleotide profile of blood cells. Cohen *et al.* [16] reported a comprehensive study on retention manipulation of bases, nucleosides and oligonucleotides using micellar solutions and metal additives. Cohen and co-workers used a conventional UV absorbance detector operating at 210 nm to study bases and nucleosides that were injected at fairly high concentrations—from 1.8×10^{-3} to 9.0×10^{-3} M. Row *et al.* [17] separated similar nucleic acid constituents using MECC and studied the effects of SDS concentration, separation voltage, injection time, injection voltage, and solute concentration on the separation.

The separation of a mixture of double-stranded DNA fragments is important for assessing the performance of polymerase chain reaction, a method of DNA amplification, or for determination restriction fragments from digests of DNA. Knowledge of the size and concentration of individual DNA fragments in a mixture is essential for manipulating these fragments, for example, for ligation. Separation of DNA fragments is traditionally done by slab gel electrophoresis. However this method is cumbersome, slow, and fairly insensitive. Several authors have studied CE methods for DNA separation [3, 18-27] to improve speed, efficiency, sensitivity and quantitation.

In this chapter, micellar electrokinetic capillary chromatography (MECC) and thermo-optical absorbance detection are applied to the analyses of a mixture of DNA fragments and a mixture of purine bases, pyrimidine bases and nucleosides.

5.2 Experimental

Bases and Nucleosides

Stock solutions of the bases adenine, cytosine, thymine and uracil, and of the nucleosides adenosine, cytidine, thymidine and uridine (all purchased from Sigma except uracil, from Eastman Kodak) were made by dissolving 5 mg each in 1.5 mL of 50% acetonitrile, 50% 5 mM sodium phosphate/1.25 mM sodium tetraborate buffer, pH 7. A stock solution of the nucleoside guanosine (Sigma) was made by dissolving 5 mg in 1.5 mL of 33% glacial acetic acid/ 67% 5 mM sodium phosphate/ 1.25 mM sodium tetraborate buffer (pH 7) then mixing in an ultrasonic bath for several minutes. The base guanine was not included because it was too difficult to dissolve in any solvent compatible with CE determination. A mixture of the nine bases and nucleosides was made by adding approximately 30 μL of each component in a microcentrifuge vial and making up a total volume of 1.74 mL with buffer: 45 mM SDS, 10 mM sodium phosphate, 2.5 mM sodium tetraborate, pH 7. The concentration of each component ranged from 2.3 to 11×10^{-4} M. Figure 5.2 shows the chemical structures of the bases and nucleosides used in this study.

Micellar electrokinetic capillary chromatography (MECC) was performed in a 42-cm-long, 50-μm-ID, 190-μm-OD fused silica capillary (Polymicro Technologies). Detection was carried out 37 cm from the injection end using the thermo-optical absorbance method described in Section 2.2 of Chapter 2. The separation buffer was made from 100 mM SDS, 25 mM sodium phosphate and 12.5 mM sodium tetraborate pH 7.5. Sample was injected electrokinetically for 5 seconds at 1 kV and separated at 7 kV. Data collection was performed as described in Chapter 2 but using a time constant of 1 second.

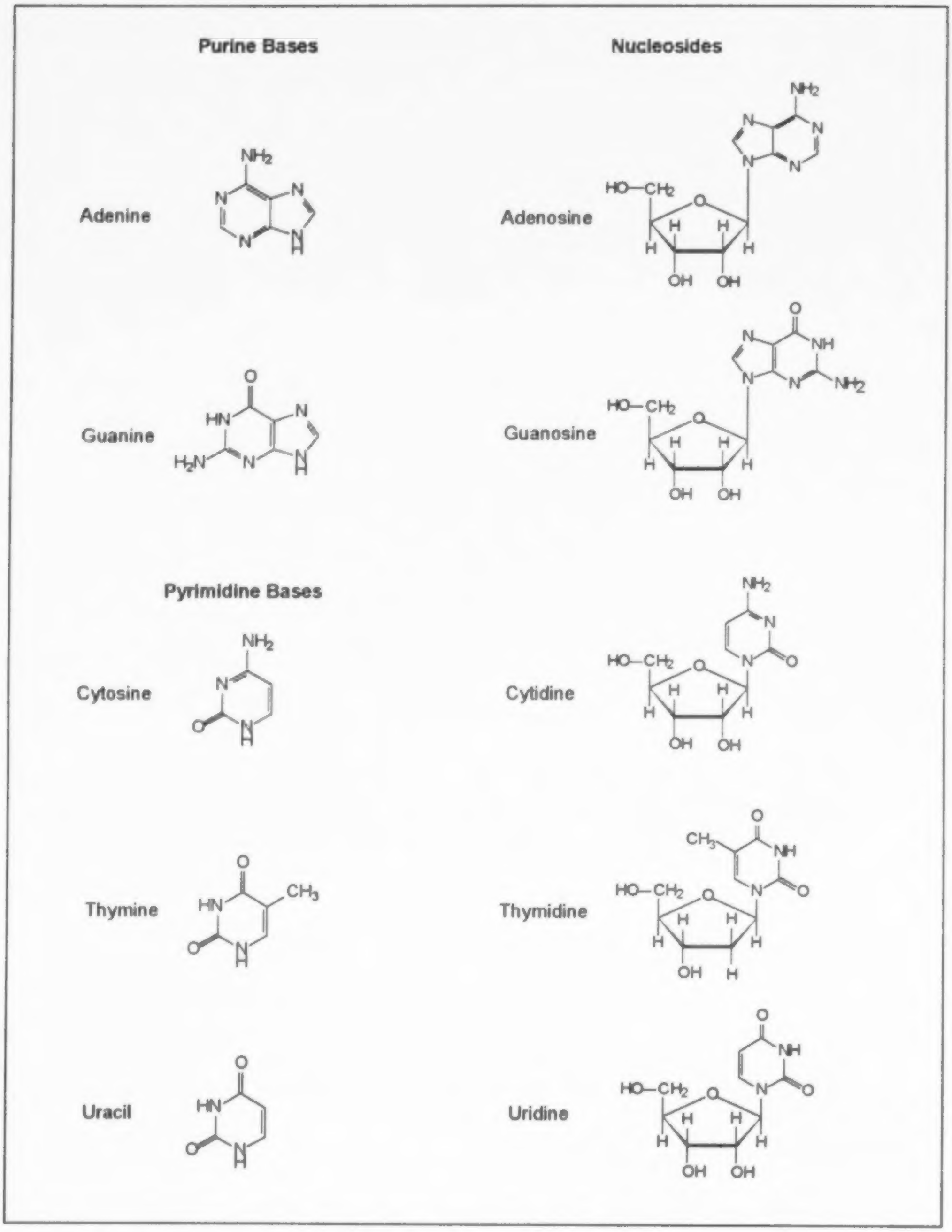


Figure 5.2 Structures of the bases and nucleosides.

Double-Stranded DNA

Two DNA mixtures were studied: a 100-base-pair ladder and a ΦX174 RF DNA/*Hae III* digest. Both of these mixtures consist of more than ten different fragments of double-stranded DNA, ranging from 72 to 2,072 base pairs (bp). Figure 5.3 shows the structure of a 3-base-pair piece of DNA.

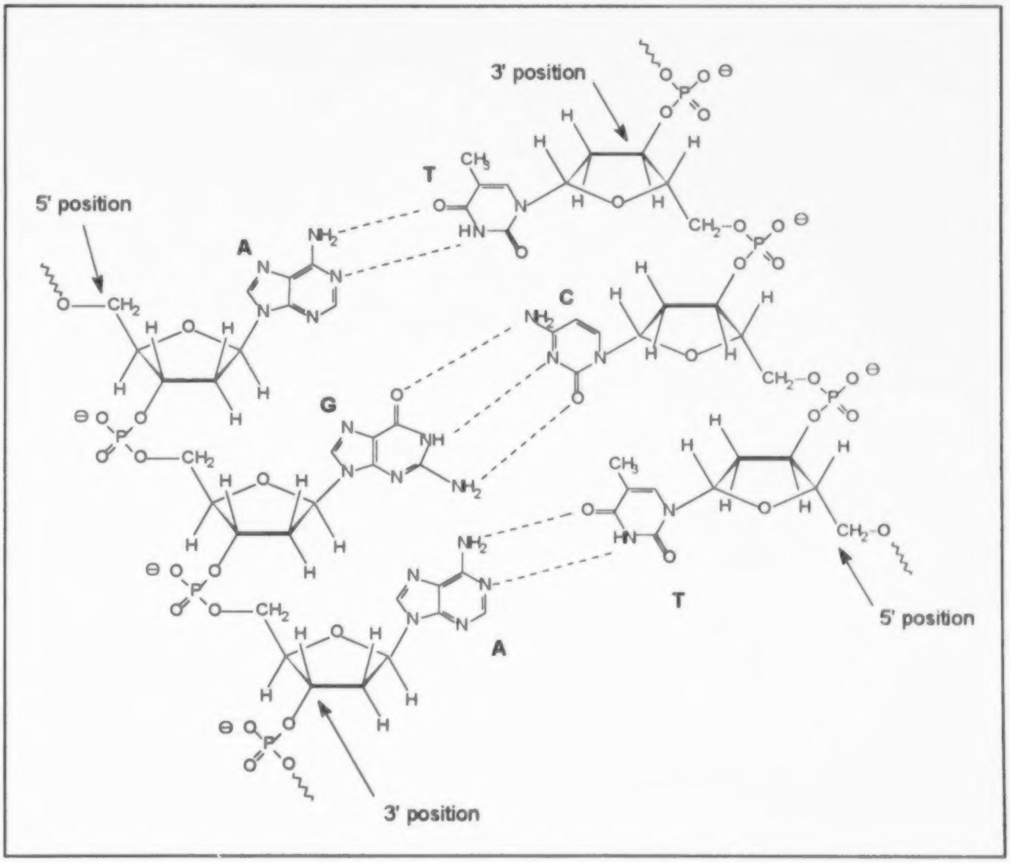


Figure 5.3 Double-stranded DNA with 3 base pairs shown.

Both DNA samples were obtained from GIBCO BRL. Only the 100-bp ladder was further diluted, 1:4 with purified water. The 100-bp ladder concentration was

250 μg/mL and the ΦX174 RF DNA concentration was 626 μg/mL. Separation of each mixture was performed in a 51-μm-ID, 360-μm-OD, 49-cm-long fused silica capillary coated inside with a 0.05-μm-thick film of methyl silicone (DB-1, J&W Scientific). The separation buffer consisted of 56 mM sodium tetraborate/boric acid, pH 8.2 to which enough hydroxypropylmethyl cellulose (HPMC) was added to make a 0.4% solution by weight. The solution was shaken vigorously for several minutes to dissolve the HPMC uniformly. Injection was done electrokinetically at 10 kV for either 5 or 16 s. Separation was run at 7.5 and 10 kV, from negative to positive high voltage (opposite to the MECC experiments). Detection was done on-column, 44 cm from the injection end, using the laser-based thermo-optical absorbance detector described in Chapter 3. A neutral density filter with optical density 0.7 was used to attenuate the UV excimer pump beam.

5.3 Results and Discussion

Bases and Nucleosides

The electropherogram of a mixture of nine bases and nucleosides separated by MECC with thermo-optical absorbance detection is shown in Figure 5.4. The separation is performed at a relatively low applied voltage of 7 kV due to the excessive heating that occurs during the separation. A current of 25 μA is observed, which is high because the ionic strength of the buffer is so high. Combined with localized heating at the detection region from the UV excimer laser, Joule heating from the separation tends to cause bubbles at voltages higher than 7 kV. The low separation voltage and slow time constant used for data collection ($\tau = 1$ s) contribute to rather poor efficiency and resolution for this separation. Meyer [28] shows that the maximum allowable time constant, τ, can be determined by Equation 5.1:

$$\tau = \theta \frac{t_m}{\sqrt{N}} \tag{5.1}$$

where θ^2 is the fraction of peak broadening, t_m is analyte migration time and N is the number of theoretical plates. The theoretical plate counts range from 41,500 for uracil to 102,000 for adenosine so, if 5% peak broadening is acceptable, then a 0.5 s time constant should have been used.

The purine and pyrimidine bases are uncharged at pH 7.5 and therefore cannot be separated by free zone capillary electrophoresis. However, micellar solutions provide the necessary partitioning effect for separating neutral species with similar structures. Cohen *et al.* [16] demonstrated the separation of bases and nucleosides using SDS micelles in a phosphate/borate buffer. They used a very high ionic strength buffer—300 mM SDS, 50 mM sodium phosphate, 25 mM sodium tetraborate—probably because the injected sample concentrations were 2×10^{-3} to 9×10^{-3} M. Cohen and co-workers showed that the solute capacity factor, k', decreased with increasing current and therefore optimized their separation of bases and nucleosides at 5.9 kV. The separation by Cohen *et al.* took 34 minutes to complete, compared with 11 minutes for the data shown in Figure 5.4 and although values were not provided their peak efficiencies looked very poor as peak widths were on the order of 15 s. Also, no indication of detection limits were given.

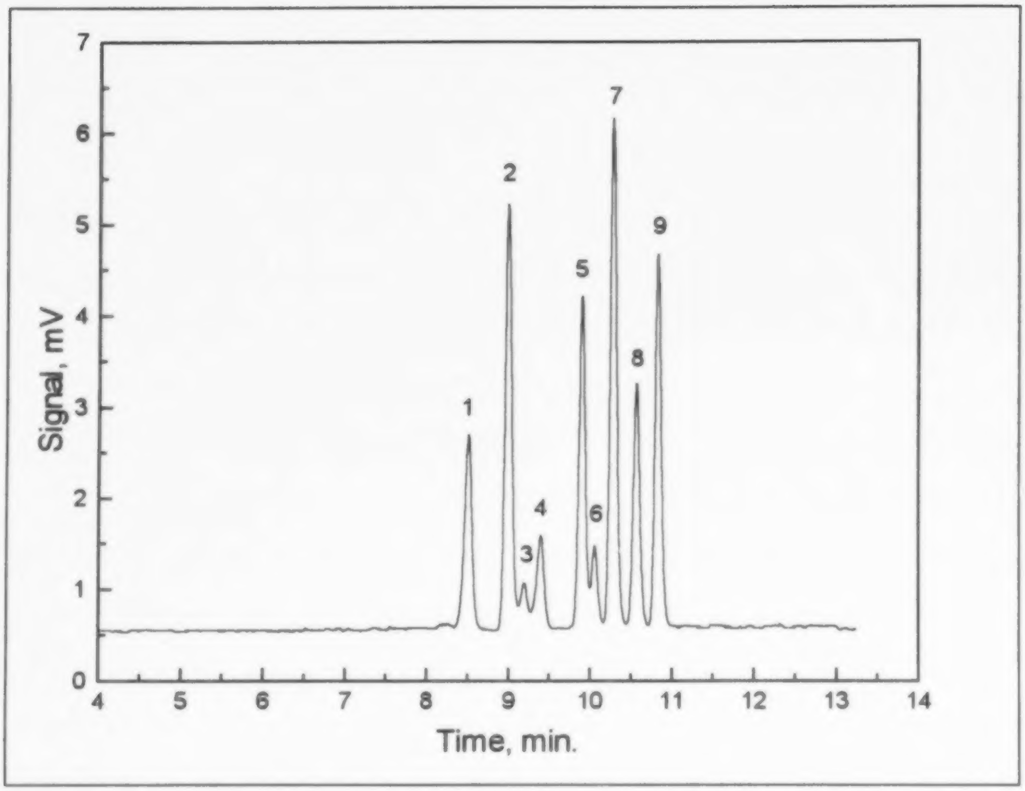


Figure 5.4 Electropherogram of a mixture of nine bases and nucleosides. A 50-μm-ID, 42-cm-long capillary was used for the separation that proceeded at 7 kV in a 100 mM SDS, 25 mM sodium phosphate, 12.5 mM sodium tetraborate, pH 7.5 buffer. Injection was done at 1 kV for 5 s and detection was performed 37 cm from the injection end. The peaks are labeled as follows, (1) uracil, (2) cytosine, (3) thymidine, (4) thymine, (5) cytidine, (6) uridine, (7) adenine, (8) guanosine, (9) adenosine. Injected concentrations of each component were approximately 5×10^{-4} M.

Row *et al.* [17] studied MECC separations of several nucleic acid constituents, mostly the 2'-deoxynucleosides. They performed their separations in buffer consisting of 75 mM SDS, 10 mM sodium phosphate, 6 mM sodium tetraborate and ran the voltage at 10 kV. The separation took 40 minutes, although a 68.5-cm-long capillary was used contributing to the long analysis time. Row and co-workers optimized the separation and achieved theoretical plate counts as high as 150,000. Detection was done with a conventional on-column UV absorbance detector operating at 256 nm. They note that the sensitivity of MECC with on-column UV detection is influenced by analyte extinction coefficients (molar absorptivity), separation efficiency, column diameter, column alignment in the detector, and characteristics of the detector itself.

The on-column thermo-optical absorbance detector has similar limitations when compared to conventional absorbance detectors, with the exception of column diameter. For the thermo-optical detector, molar absorptivity is very important because detection can only be performed at 248 nm, the UV excimer waveguide laser output wavelength. Figure 5.5 shows the UV absorbance spectrum of the purine base adenine. The molar absorptivity at 248 nm (7,700 Lmol^{-1}cm^{-1}) is only 64% of the maximum at 262 nm (12,000 Lmol^{-1}cm^{-1}) but sensitivity is sufficiently high to give good detection limits.

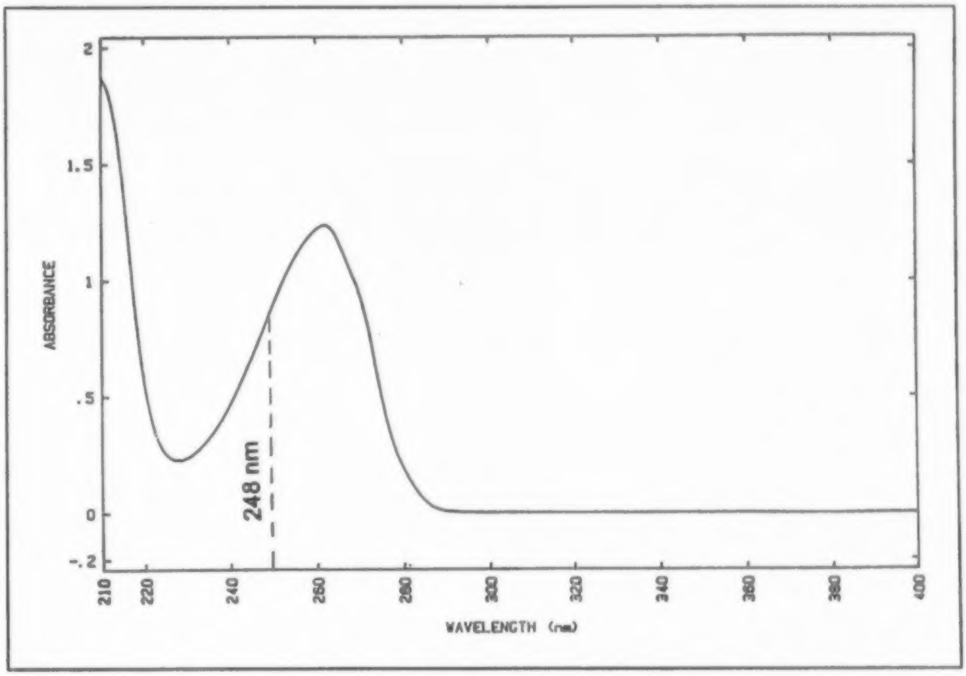


Figure 5.5 Absorbance spectrum of adenine with reference to the separation buffer. Molar absorptivity at operating wavelength, ε_{248}=7,700 Lmol^{-1}cm^{-1}, is 64% of maximum, ε_{262}=12,000 Lmol^{-1}cm^{-1}. Spectrum was obtained on a Hewlett Packard diode array spectrophotometer.

The detection limit can be calculated by two methods, similar to what was done in Chapter 2. The first way is by construction of a calibration curve to determine the sensitivity (slope). The detection limit is then calculated as 3σ/slope, where σ is the standard deviation of the base line noise. Since sensitivity varies among molecules with different molar absorptivities, a calibration curve would have to be constructed for all components in a mixture. The second way is to use Knoll's method [29] and estimate detection limits based on the maximum deviation from the mean base line signal measured over a selected time period as described by Equation 2.1 in Chapter 2. Both methods are used here. Figure 5.6 shows the calibration curve for adenine, one of the purine bases found in both DNA and RNA. Conditions for obtaining each point were the same as for the data presented in Figure 5.4.

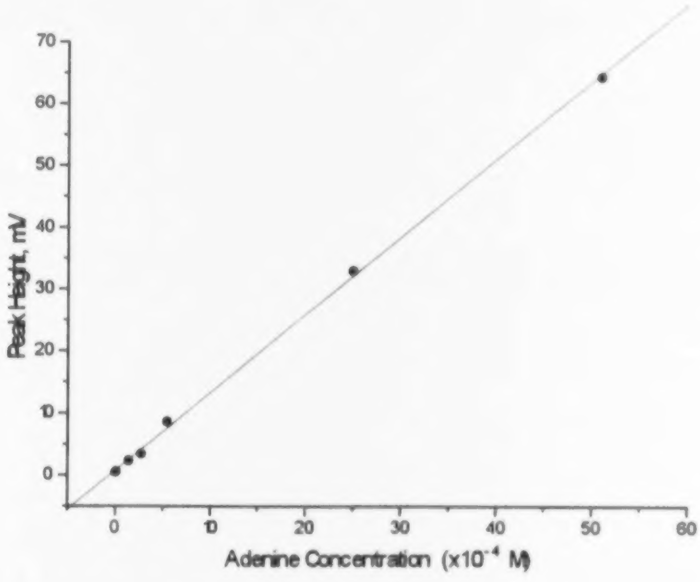


Figure 5.6 Calibration curve for adenine.

The slope of the calibration curve is 1.25×10^4 mV/M giving a concentration limit of detection, based on 3σ/slope, of 2×10^{-6} M. Injected amounts are on the order of 1 nL giving a mass detection limit of 2 fmol. Row *et al.* [17] mention that 3×10^{-5} M is approximately the concentration detection limit for their system and they reported mass detection limits for deoxynucleosides at the 50 fmol level for a signal-to-noise ratio of 4. These detection limits are very good for a conventional on-column absorbance detector but are a factor of 25 poorer than the values presented here. The influence of column alignment on detection sensitivity becomes apparent in comparing the detection limit of adenine to PTH-glycine (4×10^{-7} M, Figure 2.9, Chapter 2). The detection limit for adenine is one order of magnitude worse even though adenine has a higher molar absorptivity at λ=248 nm than PTH-glycine. The differences can be attributed mostly to detector alignment and partly to separation efficiency, as noted by Row *et al.* [17].

The detection limits for the separation of bases and nucleosides are also calculated by Knoll's method [29] and presented in Table 5.1. The volume injected is calculated using Equation 5.2, derived for electrokinetic injection in capillary zone electrophoresis [30]:

$$\mathrm{Vol}_i = \frac{(\mu_{ep} + \mu_{eo})\pi r^2 V_i t_i}{L} \tag{5.2}$$

where V_i is the injection voltage, t_i is the injection time and L is the capillary length to the detector. The calculated volume is really only an estimate because the separation is performed by MECC, not CZE. The amount injected is corrected for dilution that occurs during separation, using Equation 5.3 derived by Meyer [28] for liquid chromatography.

$$C_{det} = C_i \frac{Vol_i}{Vol_m} \sqrt{\frac{N}{2\pi}} \qquad (5.3)$$

where C_{det} is the concentration of analyte at the peak maximum, C_i is the injected concentration, Vol_m is the migration volume and N is the number of theoretical plates. The detection limit for adenine in Table 5.1 is about one order of magnitude better than that obtained from the calibration curve method. The difference is much greater for adenine than PTH-glycine (Chapter 2) and is attributed to different alignments of the detector; the calibration curve study was performed a week after the separation shown in Figure 5.4.

Table 5.1 Calculation of detection limits for bases and nucleosides.

Bases	Moles injected, fmol	Volume injected, nL	Detection limit, fmol	Detection limit, µM
Adenine	53	0.8	0.4	0.5
Cytosine	115	1.0	1.0	1.1
Thymine	53	0.9	2.2	2.3
Uracil	127	1.0	2.5	2.4
Nucleosides				
Adenosine	28	0.8	0.3	0.4
Guanosine	25	0.8	0.4	0.5
Cytidine	60	0.9	0.7	0.8
Thymidine	24	0.9	1.9	2.0
Uridine	51	0.9	2.3	2.6

The linear dynamic range of the MECC/thermo-optical detector system was evaluated for adenine, an analyte representative of the bases and nucleosides. Figures 5.7A and B show plots of the linear range of concentration versus peak height and concentration versus peak area respectively. The dynamic range of peak height (shown in Figure 5.7A) extends from 1.1×10^{-5} M to 9×10^{-3} M, the region over which the slope is 0.80. This represents a factor of 800 in analyte concentration. At higher concentrations, slight tailing is observed as analyte and separation buffer concentrations become similar. The dynamic range in peak area (shown in Figure 5.7B) extends from 1.1×10^{-5} M to 1.2×10^{-2} M and the slope of the curve is 0.98. The peak area is linear over a factor of 3,000 in analyte concentration, which is almost four times the range of peak height versus concentration.

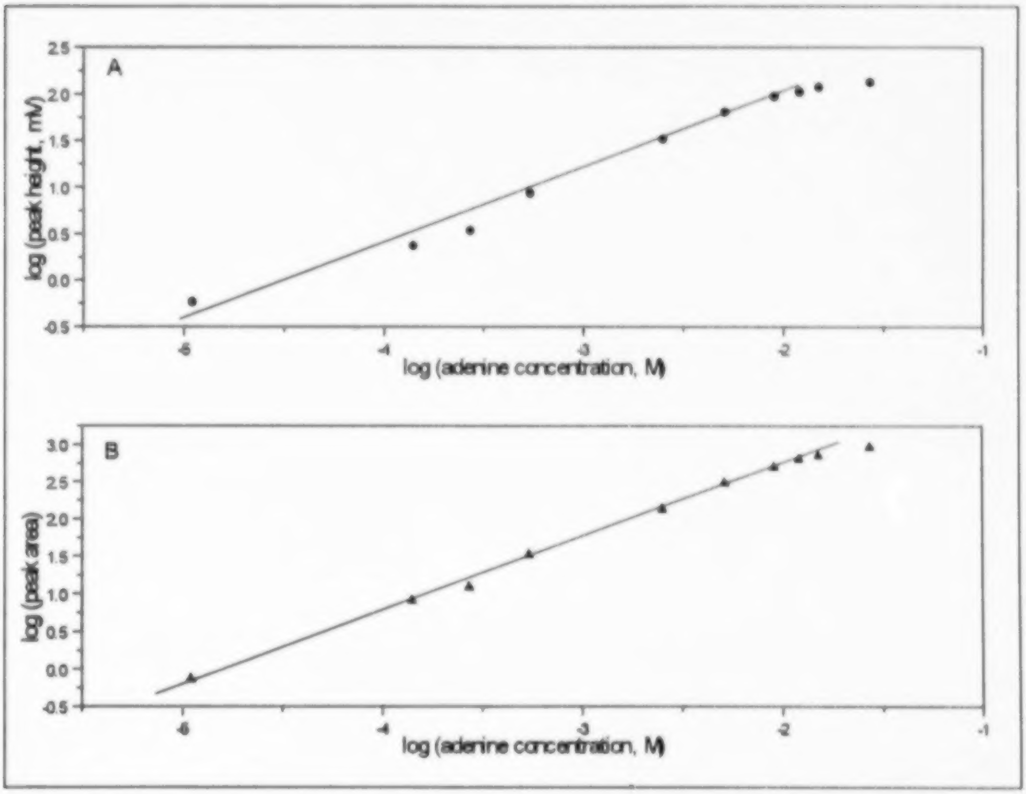


Figure 5.7 Plot of linear dynamic range of adenine for A: peak height and B: peak area. Separation conditions were the same as for data shown in Figure 5.4.

Double-Stranded DNA

DNA base pair (bp) ladders and restriction fragments are used as molecular size standards for sizing double-stranded DNA. These mixtures consist of multiple repeats of a nucleic acid sequence to allow accurate determination of molecular sizes. Two mixtures have been used in the work described in this chapter. The first is a 100-bp ladder consisting of 15 different sizes of DNA ranging in length from 100 to 1,500 bp, in increments of 100. It is prepared from a plasmid and therefore the mixture contains the vector-DNA, 2,072 bp in length, originating from the plasmid. The second is a restriction fragment mixture prepared from purified ΦX174 RF DNA that has been digested to completion with *Hae III*, a restriction endonuclease. Bacteriophage ΦX174 is a small, icosahedral virus that infects *E. coli*. RF refers to the replicative form of the DNA, which is a double stranded circle. *Hae III* cuts the circular DNA in 11 locations, generating 11 different sizes, or fragments, ranging from 72 to 1,353 base pairs. No vector piece is present since the DNA template (ΦX174 RF DNA) is completely digested.

Typically, the standards are run in a lane parallel to the unknown DNA of interest during slab gel electrophoresis. Figure 5.8 shows copies of slab gel electropherograms obtained from the supplier literature for the 100-bp ladder and *Hae III* digest. The bands migrate slowly through the gel and require approximately 2 hr for separation. The bands are stained to make them visible or, if ethidium bromide was initially in the electrophoresis buffer, the bands are viewed under UV light. Ethidium bromide is an intercalating agent that fluoresces only when it is incorporated between the strands of double-stranded DNA.

Slab gel electrophoresis is a traditional and reliable way of separating DNA; however, it is time consuming and tedious. Slabs often have to be 30×20 cm or larger to provide sufficient resolution of the bands. If radiolabels are used to detect DNA, analysis takes approximately 24 hr because an audioradiogram of the gel has to be made so the bands can be visualized. In the past five years, capillary gel electrophoresis has gained popularity as a fast, efficient and highly sensitive means of separating double-stranded DNA compared to the slab gel method.

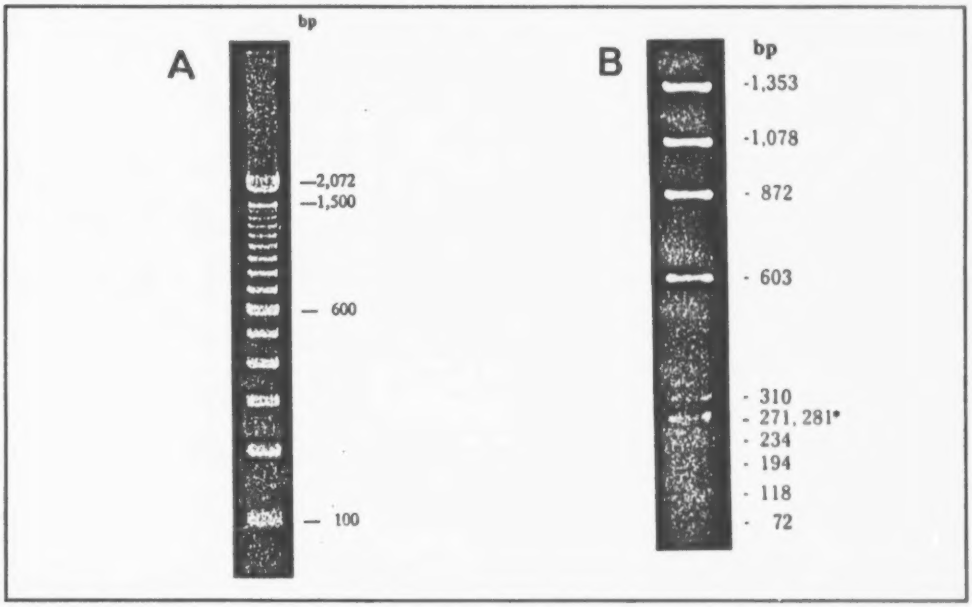


Figure 5.8 Slab gel electropherograms of molecular size markers used to size double-stranded DNA. **A**: 100-bp ladder run in 2% agarose gel and stained with ethidium bromide. The 100 and 600 bp fragments are spiked to provide easier orientation. **B**: ΦX174 RF DNA/*Hae III* restriction fragments run in 6% polyacrylamide gel and stained with ethidium bromide. *Fragments of length 271 and 281 bp tend to migrate very close together. [3]

DNA is highly negatively charged, with charge proportional to fragment length. As a result, electrophoretic mobility is almost independent of molecular size so DNA cannot simply be separated by capillary zone electrophoresis—it requires a sieving medium. Separation is based on size alone in gel electrophoresis. Polyacrylamide and agarose are the most common sieving mediums for separating DNA and have been used in capillaries for a few years now. When cross-linked, polyacrylamide can resolve pieces of single-stranded DNA that differ by one nucleotide in length [31]. Non-crosslinked (linear) polyacrylamide is considered an entangled polymer and, recently, several authors have demonstrated capillary electrophoretic separations of double-stranded DNA base pair ladders and restriction fragments using linear polyacrylamide [3, 20-22, 25] and melted (<1%) agarose [27]. Linear polyacrylamide and agarose are essentially solid gels and therefore are difficult to push out of a capillary for refilling. Usually, new capillaries must be prepared. Alternative sieving mediums are the cellulose-based entangled polymers that have recently demonstrated good electrophoretic separations of double-stranded DNA in capillaries [3, 18, 19, 23, 24, 26]. Methylcellulose, hydroxyethyl cellulose (HEC), and hydroxypropylmethyl cellulose (HPMC) are less viscous than polyacrylamide and can easily be pushed out of capillaries, which can then be reused.

---

[3] Photographs were obtained from quality control assay literature, GIBCO BRL Life Technologies, Inc.

Because DNA is anionic, electrophoretic separation is run with reversed polarity to CZE and MECC: from negative to positive. The effect of injecting at the negative end and detecting at the positive end is that electroosmosis opposes the flow of analyte. The effect of electroosmosis is not observed in agarose (2%) or cross-linked polyacrylamide separations because these are very solid gels. However, entangled polymer solutions are made in aqueous buffers and are prone to electroosmotic flow. To reduce electroosmosis, silanol groups on the capillary walls must be blocked—effectively eliminating the electric double layer responsible for electroosmotic flow. Several methods are available to reduce electroosmosis, ranging from silanization of the capillary with methacryloxypropyltrimethoxy silane [3, 32] to using gas chromatography capillaries coated with stationary phases [20, 23] as in this work. Figures 5.9 and 5.10 show the separation of double-stranded DNA by capillary gel electrophoresis in hydroxypropylmethyl cellulose (HPMC) using the thermo-optical absorbance detection system.

The electropherograms shown in Figures 5.9 and 5.10 demonstrate good resolution of the double-stranded DNA fragments. The 600-bp strand of the 100-bp ladder migrates as two bands, a behavior that has been seen in slab gels, in entangled polymers by others [3] and is noted in the supplier literature as an anomalous migration. The resolution for both samples is at the least comparable to other published work and in some cases better. Schwartz *et al.* [23] were unable to resolve the 271/281 bp fragments of ΦX174 DNA in 0.5% HPMC unless they added 10 μM ethidium bromide to the separation buffer. Ethidium bromide tends to alter the DNA structure slightly, and perhaps the charge, aiding in resolution of the 271 and 281 bp fragments. Thirty-three minutes were required for their separation. In this work, fragments 271 and 281 are not quite base line resolved but still distinguishable without addition of ethidium bromide (see Figure 5.10). The total analysis is complete within 22 minutes. A peak efficiency of 790,000 theoretical plates was calculated for the 200-bp fragment in Figure 5.9. This value is fairly close to the expected value considering that the field strength during separation was only 205 V/cm.

The maximum separation voltage that can be used with the thermo-optical absorbance detector is about 10 kV because of significant absorbance of the buffer components at 248 nm. Above 10 kV, the Joule heating in the capillary becomes too great, specifically at the detection region were there is the additive heat from the pulsed UV laser. The net effect is that bubbles are formed in the capillary at the detector, presumably due to boiling of the buffer. Most authors used HPMC or other celluloses dissolved in TBE buffer (Tris-Borate-EDTA) to separate DNA fragments. Unfortunately, EDTA absorbs strongly at the detection wavelength. Borate buffer alone (sodium tetraborate/boric acid) was found to absorb less than Tris-borate so the former

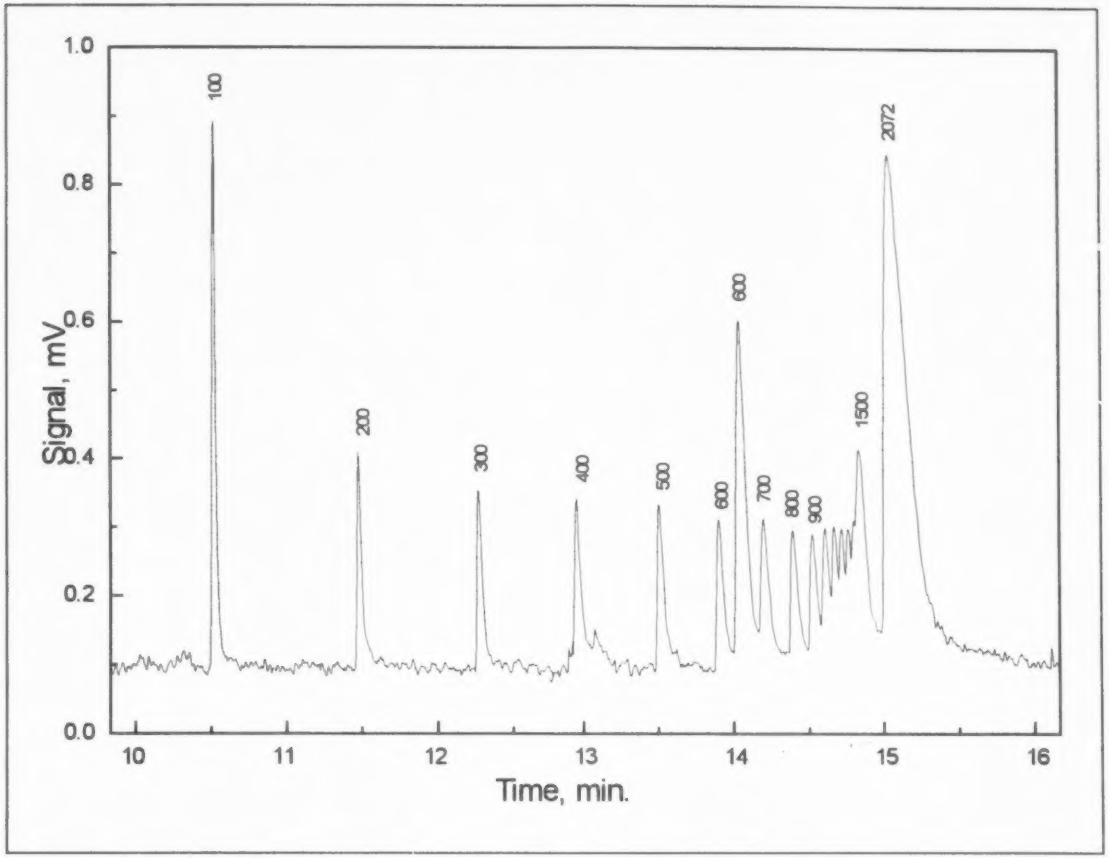


Figure 5.9 Electropherogram of the 100-bp ladder. A 49-cm long, 50-μm-ID coated (DB-1) capillary was used for the separation that proceeded at 10 kV in a 0.4% HPMC in 56 mM sodium tetraborate/boric acid, pH 8.2 buffer. Electrokinetic injection of the 250 μg/mL sample was done at 10 kV for 5s and detection performed 5 cm from the end.

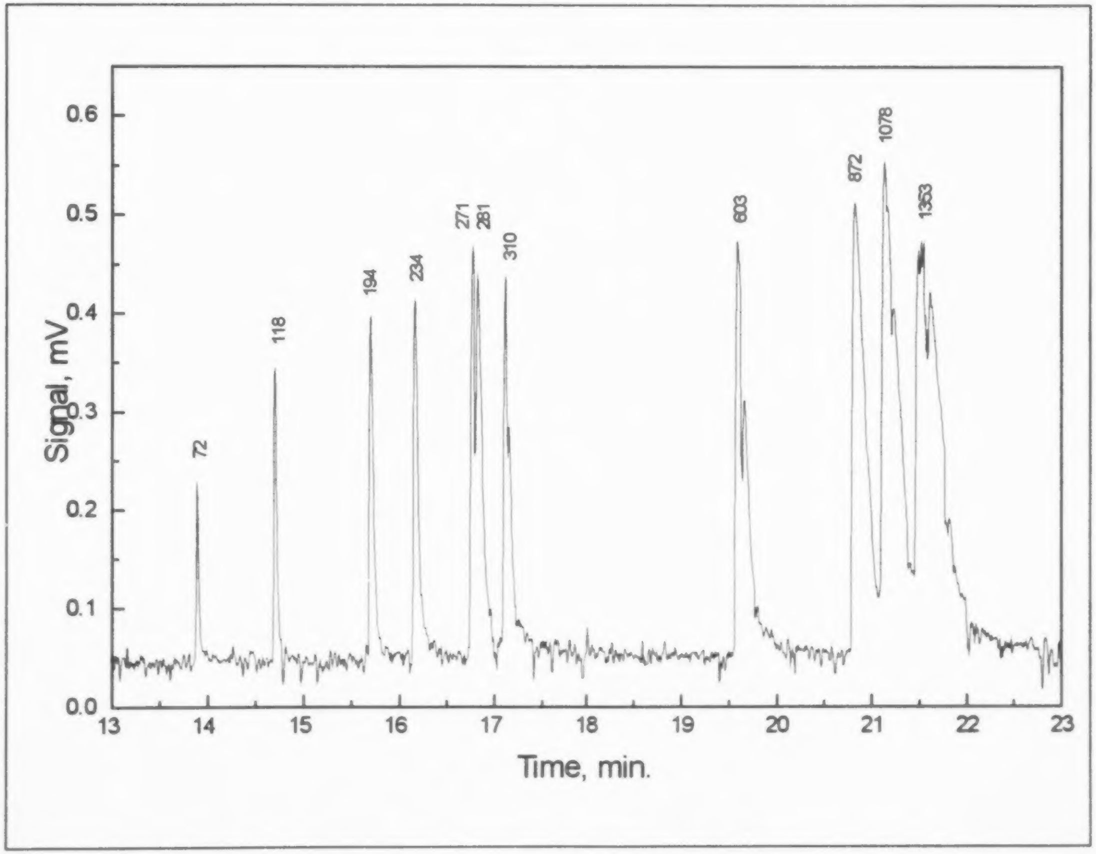


Figure 5.10 Electropherogram of the *Hae III* digest of ΦX174 RF DNA. A 49-cm long, 50-µm-ID coated (DB-1) capillary was used for the separation that proceeded at 7.5 kV in a 0.4% HPMC in 56 mM sodium tetraborate/boric acid, pH 8.2 buffer. Electrokinetic injection of the 625 µg/mL sample was done at 10 kV for 16s and detection performed 5 cm from the end.

was used for the separations seen in Figures 5.9 and 5.10. Buffer absorption and subsequent heating are also to blame for the noisy base line signals.

Double-stranded DNA has a huge molar absorptivity at $\lambda = 260$ nm— 1,000,000 $\mathrm{Lmol^{-1}cm^{-1}}$; however, direct UV absorbance detection is not used for slab gels since it is difficult to record the absorbance signal. Instead, a fluorescent dye is usually attached to or incorporated into the DNA and a photograph of the fluorescence is taken. Fluorescent detection of double stranded DNA in capillaries is the most sensitive [3, 24]. However, conventional UV absorbance methods are also used with detection at both 260 nm [18, 19, 23, 25, 27] and 254 nm [20-22, 26]. The electropherograms in Figures 5.9 and 5.10 were recorded using the UV thermo-optical absorbance detector at 248 nm, close to the absorbance maximum.

The limit of detection is difficult to estimate from the electropherograms alone because the sample concentration is for all DNA strands together. Chin and Colburn [18] indicate the sensitivity is 2 μg/mL/fragment for counter-migration capillary electrophoresis of DNA restriction fragments whereas Schwartz *et al.* [23] calculate the minimum concentration for a 118 bp fragment to be 8 ng/mL at a signal-to-noise ratio of 2. Schwartz and co-workers constructed a calibration curve using successive dilutions of a DNA stock solution and plotted peak height versus concentration for three different DNA sizes. The detection sensitivity in μg/mL is much higher for shorter strands that for longer strands of DNA within a DNA ladder or digest.

The detection limit was estimated for the electropherograms in Figures 5.9 and 5.10 using Knoll's method [29], which is equivalent to a signal-to-noise ration of 3. To estimate the concentration of the 118-bp fragment in Figure 5.10, one has to assume that *Hae III* digests the entire amount of ΦX174 DNA into 11 fragments. The molar concentration of each fragment was calculated to be 1.8×10^{-7} M, which was then converted to 14 μg/mL for the 118-bp fragment. The detection limit for the 118-bp fragment is, therefore, 1 μg/mL. This result is comparible to Chin and Colburn's result [18] but much worse than Schwartz *et al.* [23]. Absorption of the HPMC buffer in the thermo-optical detector system causes a significant background signal, reducing the ability to achieve low detection limits. The molar limit of detection, 1.8×10^{-7} M, is similar to that obtained for PTH-glycine in Chapter 2 even though the molar absorptivity of DNA is much larger than for PTH-gly. The base line noise was used to obtain two of the four variables for calculating detection limit by Knoll's method. Therefore the limiting factor is noise arising from absorption of the buffer components used in the separation.

5.4 Conclusions

The high efficiency separation ability of capillary electrophoresis and the high sensitivity detection ability of thermo-optical absorbance have been used in this chapter to analyse double-stranded DNA and its constituents. The nucleotides (nucleoside + phosphate) were not studied because they are charged species and therefore less challenging to separate than the uncharged bases and nucleosides. The CE/thermo-optical detection technique demonstrated excellent linearity in peak height and area for the bases and nucleosides. Detection limits were generally better than those obtained by traditional absorbance methods with the advantage that very small volumes of sample could be analysed.

Separation of double-stranded DNA is an important diagnostic for PCR (polymerase chain reaction). Fast separations were demonstrated in this chapter but detection limits were not improved over those obtained from traditional on-column CE detectors. Efforts must be made to reduce background noise caused by absorption of the buffer. Phosphate may prove better than borate, and different celluloses or other entangled polymers need to be investigated to find the least interfering separation medium. The main disadvantage of capillary versus slab gel electrophoresis is the inability to run multiple samples simultaneously. Since base pair ladders are used as molecular weight standards, they should be run at the same time as the unknown DNA sample. Development of multiple capillary analysis systems is necessary to overcome this disadvantage, a few examples of which have already been shown [33-35].

Separation of single-stranded DNA, specifically for DNA sequencing, would be an interesting application of CE/thermo-optical absorbance detection technique. Preliminary studies showed that on-column detection was impossible since polyacrylamide absorbs strongly at 248 nm. The gel was destroyed within minutes because the heating at the detection region boiled the gel. A post-column detector would be necessary since polyacrylamide is the only suitable medium for resolving DNA fragments that differ in length by a single nucleotide. Also, absorbance detection would require peak height encoding of the four DNA bases, a procedure that has been surpassed in effectiveness by fluorescent dye encoded methods.

The analysis techniques described in this chapter are not limited to studying DNA and could be applied to a wide range of samples. One example is the separation and detection of peptides and proteins. These molecules have been separated by CE and studied extensively by several authors but application of thermo-optical absorbance detection has yet to be shown.

5.5 References

1. S. Wu and N.J. Dovichi, *Talanta* **39**, 173-178 (1992).
2. D.Y. Chen, H.R. Harke and N.J. Dovichi, *Nucleic Acids Research* **20**, 4873-4880 (1992).
3. D. Figeys, E. Arriaga, A. Renborg and N.J. Dovichi, *J. Chromatogr.* **Submitted**, (1993).

4. J.Y. Zhao, P. Diedrich, Y. Zhang, O. Hindsgaul and N.J. Dovichi, *J. Chromatogr.* **Submitted**, (1993).

5. J.Y. Zhao, N.J. Dovichi, O. Hindsgaul, S. Gosselin and M.M. Palcic, *Science* **Submitted**, (1993).

6. B.L. Hogan and E.S. Yeung, *Anal. Chem.* **64**, 2841-2845 (1992).

7. M. Ng, T.F. Blaschke, A.A. Arias and R.N. Zare, *Anal. Chem.* **64**, 1682-1684 (1992).

8. R.E. Milofsky and E.S. Yeung, *Anal. Chem.* **65**, 153-7 (1993).

9. A.E. Bruno, A. Paulus and D.J. Bornhop, *Appl. Spectrosc.* **45**, 462-467 (1991).

10. M. Yu and N.J. Dovichi, *Mikrochim. Acta* **111**, 27-40 (1988).

11. M. Yu and N.J. Dovichi, *Anal. Chem.* **61**, 37-40 (1989).

12. M. Yu and N.J. Dovichi, *Appl. Spect.* **43**, 196-201 (1989).

13. K.C. Waldron, S. Wu, C.W. Earle, H.R. Harke and N.J. Dovichi, *Electrophoresis* **11**, 777-780 (1990).

14. K.C. Waldron, J. Miller and N.J. Dovichi, *Analytical Sciences* **7**, 959-961 (1991).

15. K.C. Waldron and N.J. Dovichi, *Anal. Chem.* **64**, 1396-1399 (1992).

16. A.S. Cohen, S. Terabe, J.A. Smith and B.L. Karger, *Anal. Chem.* **59**, 1021-1027 (1987).

17. K.H. Row, W.H. Griest and M.P. Maskarinec, *J. Chromatogr.* **409**, 193-203 (1987).

18. A.M. Chin and J.C. Colburn, *Am. Biotechnol. Lab.* **7**, 16 (1989).

19. P.D. Grossman and D.S. Soane, *J. Chromatogr.* **559**, 257-277 (1991).

20. A. Guttman and N. Cooke, *J. Chromatogr.* **559**, 285-294 (1991).

21. A. Guttman and N. Cooke, *Anal. Chem.* **63**, 2038-2042 (1991).

22. A. Guttman, B. Wanders and N. Cooke, *Anal. Chem.* **64**, 2348-2351 (1992).

23. H.E. Schwartz, K. Ulfelder, F.J. Sunzeri, M.P. Busch and R.G. Brownlee, *J. Chromatogr.* **559**, 267-283 (1991).

24. H.E. Schwartz and K.J. Ulfelder, *Anal. Chem.* **64**, 1737-1740 (1992).

25. A. Paulus and D. Husken, *Electrophoresis* **14**, 27-35 (1993).

26. M. Strege and A. Lagu, *Anal. Chem.* **63**, 1233-1236 (1991).

27. P. Bocek and A. Chrambach, *Electrohporesis* **12**, 1059-1061 (1991).

28. V.R. Meyer, *J. Chromatogr.* **334**, 197-209 (1985).

29. J.E. Knoll, *J. Chrom. Sci.* **23**, 422-425 (1985).

30. R.A. Wallingford and A.G. Ewing, *Adv. Chromatogr.* **29**, 1-76 (1989).

31. M.J. Rocheleau and N.J. Dovichi, *J. Microcol. Sep.* **4**, 449-453 (1992).

32. S. Hjerten, *Electrophoresis* **10**, 23-29 (1989).

33. J.Z. Zhang, *Personal Communication* (1993).

34. H. Kambara and S. Takahashi, *Nature* **361**, 565-566 (1993).

35. X.C. Huang, M.A. Quesada and R.A. Mathies, *Anal. Chem.* **64**, 2149-2154 (1992).

CHAPTER 6

CONCLUSIONS
AND
FUTURE DIRECTIONS

6.1 Discussion

Introduction

Analysis of biomolecules requires ever increasing sophistication in analytical techniques. Because many samples are present only in very small quantities, the role of high sensitivity analysis of biomolecules has become increasingly important in biology, medical research, and the biotechnology and pharmaceutical industries. Development and application of micro-scale analytical techniques such as capillary electrophoresis (CE), thermo-optical absorbance detection, and miniaturized peptide sequencing were presented in this thesis in an effort to address a few of the challenging goals of biochemical research.

CE is a relatively new separation technique well suited to the determination of minute quantities of biomolecules. High sensitivity analysis of such minute quantities of sample often requires the use of lasers in analytical chemistry. Laser-based methods of detection take advantage of the spatial coherence of lasers and the ability to tightly focus a laser beam in order to probe the extremely small volumes encountered in CE. A detailed description of CE and the thermo-optical absorbance detector were presented in Chapter 1.

In Chapter 2 of this thesis, the fast and efficient separation ability of CE was combined with the high sensitivity detection ability of laser-based thermo-optical absorbance detection for the determination of derivatized amino acids. Amino acids are important biomolecules because they are the building blocks of peptides and proteins, which constitute over eighty percent of the organic compounds in living organisms. The derivatized amino acids studied in Chapter 2 were of particular importance to peptide and protein sequencing: the process of determining of the order in which amino acids are linked.

Peptide/protein sequencing is an integral part of understanding peptide/protein structure and function, and accessing the genes that code for peptides and proteins. The basic principles behind sequencing and the Edman method of stepwise polypeptide degradation were discussed in Chapter 1, section 1.4. Chapter 3 presented details of manual Edman Degradation for peptide sequencing, particularly as a method of assessing the ability of CE/thermo-optical detection to identify the Edman degradation products: phenylthiohydantoin amino acids.

In Chapter 4, the results and implications of work presented in the first three chapters were brought together to design a miniaturized peptide sequencer. Details on the construction and characteristics were presented and compared to commercially available systems and to similar work by other authors. One cycle of Edman degradation was shown to illustrate the potential of the miniaturized instrument to sequence femtomole amounts of peptide. Further developments are underway to improve the sequencer that was described in Chapter 4. Some of these efforts and ideas are presented in the next section of this chapter.

Chapter 5 presented two more applications of capillary electrophoresis and laser-based thermo-optical absorbance detection for analysis of minute quantities of biomolecules. First, micellar electrokinetic capillary chromatography (MECC) was used to show the separation of purine and pyrimidine bases and nucleosides. Second, gel electrophoresis in entangled polymers was used to show the separation of double stranded DNA fragments. The results presented in Chapters 2 and 5 are just a hint of the ability of CE and thermo-optical detection to become routine micro-scale analytical techniques for separation and detection of many different samples.

Discussion

Capillary electrophoresis is considered to be a complimentary separation method to high performance liquid chromatography (HPLC). The results in Chapters 2, 3 and 4 suggest that, for the separation of phenylthiohydantoin (PTH) amino acids, CE could replace the routine use of HPLC. Absorbance based detection by the thermo-optical method was shown, in Chapter 2, to provide superior concentration and mass detection limits compared to other on-column absorbance methods in CE. Sub-femtomole detection limits of derivatized amino acids represents a 1,000-fold improvement in detection capability. Ten minute separations represents a 3-fold improvement in analysis time compared to HPLC. Similar improvements in separation efficiency and detection sensitivity between slab and capillary gel electrophoresis were shown in Chapter 5.

The results from Chapter 4 (miniaturized peptide sequencer) are encouraging and, while some improvements should be considered, the modifications and applications of this technology are many. The glass fibre filter used in the reaction chamber was precycled with Polybrene using a commercial sequencer [1] however precycled mats could be made using the procedure described by Touchstone and co-workers [2]. The method of inserting the glass fibre filter into the reaction chamber could be used to insert electroblotted peptides. It would be simple to cut mats out of peptide-blotted PVDF membranes, using the wide capillary, and then assemble the reaction chamber with epoxy. The sample preparation time would be similar to cutting the membrane up into little strips and placing them into the ABI reaction cartridge [3]. Another possibility for sample loading is to pack a capillary with chromatography stationary phase that would retain the peptide during coupling and cleavage steps. Other authors have used similar methods for gas-phase [4] and solid-phase [5] sequencing. The packed capillary could become the upper part of the reaction chamber or simply be inserted into the pre-assembled reaction chamber. A similar procedure is to derivatize the reaction chamber walls with Polybrene or another peptide binding substance and then let the peptide sample adsorb onto the walls, a method similar to that described by Tarr [6].

The miniaturized peptide sequencer would be very well suited to sequencing peptides from peptide libraries. Synthetic peptide combinatorial libraries (SPCL) are complex mixtures of peptides generated as tools for basic research and drug discovery [7]. The peptides are synthesized from the C-terminal end by attaching the first amino acid to a resin bead then building out to the N-terminus. The peptide on a single bead is sequenced after it is found to be biologically active. The beads are very small (~100 μm diameter) and tend to get lost in the large volume of the ABI reaction cartridge. Often, coupling and cleavage reagents during Edman degradation flow past the bead without ever touching it [8]. The miniaturized peptide sequencer has a much smaller reaction chamber and would, therefore, be ideal for sequencing peptide attached to a single bead. SPCLs are used extensively in immunoassays, enzyme inhibition assays and to study peptide reactions and interactions.

While thermo-optical absorbance detection has shown excellent sensitivity for detecting PTH amino acids, it cannot compare with the exquisite sensitivity of laser-induced fluorescence detection. Separation and detection of fluorescien thiohydantoin (FTH) amino acids using CE has been already demonstrated by Wu and Dovichi [9]. A logical extension of the work in Chapter 4 is to perform Edman degradation in the miniaturized sequencer using fluorescein isothiocyanate (FITC) and then determine the FTH amino acid products with CE and laser-induced fluorescence. A few modifications of the sequencer would have to be made since double coupling is required. FITC does not react quantitatively with peptides and proteins so double coupling with PITC or MITC (methylisothiocyanate) is required. This chemical procedure has been described in several papers [10-15]; however, high sensitivity detection of small amounts of the FTHs was not reported until Wu's study [9].

Changes to the miniaturized sequencer design would entail adding extra inputs to the first valve and reagent vessels to distribute the FITC sequencing reagent and wash solvents. In-line conversion of the ATZ amino acid product to the thiohydantoin can be accomplished in a conversion chamber of similar configuration to the reaction chamber. Reagent delivery to the conversion chamber can be done via a second valve to prevent exposing the peptide to aqueous acid. Addition of an electrophoresis separation capillary using a miniature zero dead volume "T" will allow determination of the total amino acid product formed from degradation. Figure 6.1 shows a schematic of a peptide sequencer with in-line conversion and on-column separation/detection of

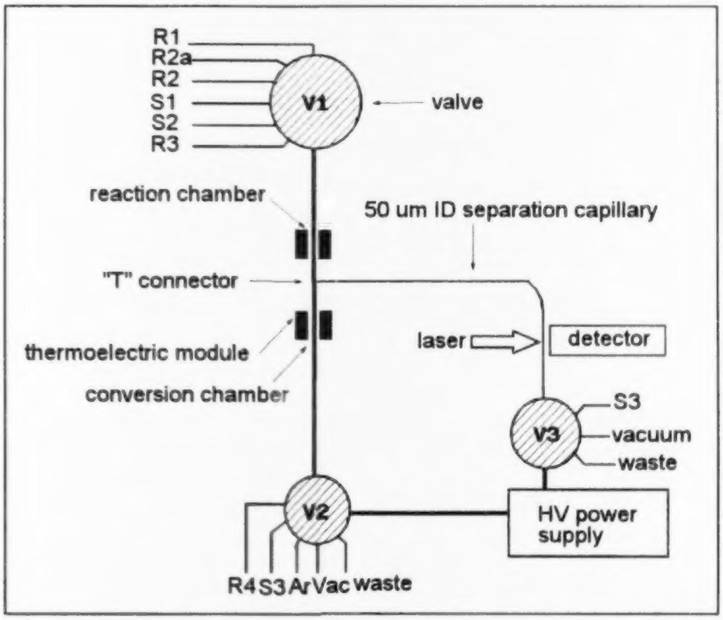


Figure 6.1 Attomole peptide sequencer schematic. R1-R4 represent the reagents required for FITC/PITC degradation, S1 and S2 represent the solvents required during sequencing, S3 represents the CE separation buffer, Vac represents vacuum-assisted solvent removal.

the sequencing products. On-column separation and laser-based fluorescence detection should provide attomole level or better peptide sequencing.

The miniaturized peptide sequencer is also a prime candidate for interface with a mass spectrometer (MS) for detection of Edman degradation products. Many people have studied CE-MS methods of separation and high sensitivity detection as discussed in a recent review article by [16]. The two main aspects of interfacing CE with MS are (1) compatibility of the high voltage associated with CE separation and, (2) compatibility of the minute sample volume eluted after separation. The second aspect is of importance to connecting the miniaturized sequencer directly to a mass spectrometer. Building an interface should be fairly straightforward since much work has already been done for CE-MS. Determination of the derivatized amino acids can be performed *without* separation since MS gives a unique spectrum for almost all the physiological amino acids. The exception is leucine and isoleucine, which have the same fragment mass. To improve detection, Edman based reagents with enhanced mass spectral sensitivity are being developed [17, 18] for sequencing by the automated gas- or solid-phase method. Direct determination of the cleaved amino acid is done by MS. Since mass spectrometry requires injection of very small volumes, the miniaturized peptide sequencer would be well suited to deliver concentrated amino acid product in microlitre volumes. Figure 6.2 shows a schematic of a miniaturized peptide sequencer with mass spectral identification of the thiohydantoin amino acid products from degradation. This configuration should provide both high sensitivity and speed for peptide and protein sequencing.

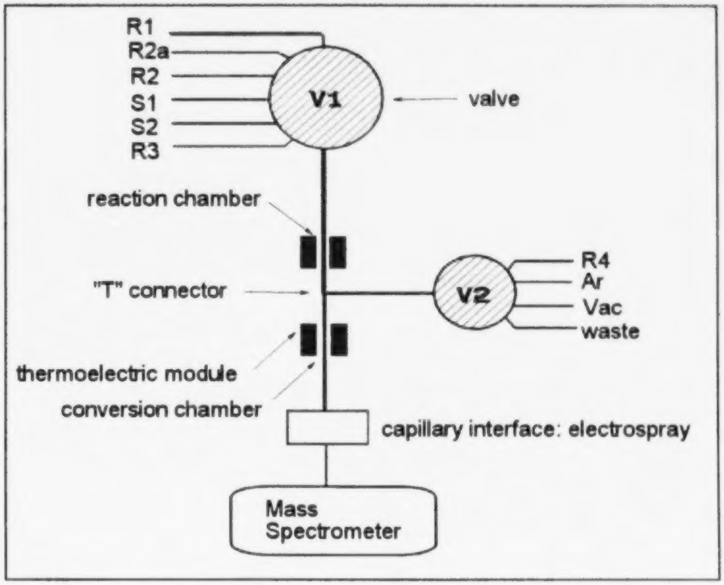


Figure 6.2 Schematic of a miniaturized peptide sequencer with mass spectral identification of amino acid derivatives. R1-S2 represent the reagents and solvents required for Edman degradation.

Commercialization of the micro-scale analytical techniques described in this thesis will be possible in the future. This step would mean genuine acceptance and applicability of the analytical methods. Analysis of samples at lower and lower levels of detection will always be a challenge. Capillary electrophoresis, laser-based thermo-optical absorbance detection and miniaturized peptide sequencing help to meet this challenge. The results presented in this thesis will hopefully mean long term advances in routine analysis in the biomedical fields.

6.2 References

1. T. Bures, *Personal Communication* (1992).
2. B. Touchstone, Z.-W. Gu and C.-Y. Yang in *Methods in Protein Sequence Analysis, Proceedings of the 7th International Conference, Berlin, July 3-8, 1988,* (eds. B. Wittmann-Liebold) 122-128 (Springer-Verlag; Berlin, 1989).

3. M.R. Carpenter, *Personal Communication* (1992).

4. J.E. Shively, P. Miller and M. Ronk, *Anal. Biochem.* **163**, 517-529 (1987).

5. S.-P. Liang and R.A. Laursen, *Anal. Biochem.* **188**, 366-373 (1990).

6. G.E. Tarr, *J. Prot. Chem.* **7**, 293-295 (1988).

7. R.A. Houghten, C. Pinilla, J.R. Appel, C.T. Dooley, P. Blanc and R.R. Tuttle, Abstract #TH1 in *Thirteenth American Peptide Symposium*. Edmonton, Alberta, June 20-25, 1993.

8. N. Sepetov, *Personal Communication* (1993).

9. S. Wu and N.J. Dovichi, *Talanta* **39**, 173-178 (1992).

10. H. Kawauchi, K. Tuzimura, H. Maeda and N. Ishida, *J. Biochem.* **66**, 783-789 (1969).

11. H. Kawauchi, K. Muramoto and J. Ramachandran, *Int. J. Peptide Protein Res.* **12**, 318-324 (1978).

12. H. Maeda, N. Ishida, H. Kawauchi and K. Tuzimura, *J. Biochem.* **65**, 777-783 (1969).

13. K. Muramoto, H. Kawauchi and K. Tuzimura, *Agric. Biol. Chem.* **42**, 1559-1563 (1978).

14. K. Muramoto, H. Kamiya and H. Kawauchi, *Anal. Biochem.* **141**, 446-450 (1984).

15. B. Wittmann-Liebold, J. Shan-Wei and J. Salnikow in *Protein/Peptide Sequence Analysis: Current Methodologies,* (eds. A.S. Bhown) 119-134 (CRC Press; Boca Raton, 1988).

16. W.M.A. Niessen, U.R. Tjaden and J. vanderGreef, *J. Chromatogr.* **636**, 3-19 (1993).

17. R. Aebersold, E.J. Bures, M. Namchuck, M.H. Goghari, B. Shushan and T.C. Covey, *Protein Science.* **1**, 494-503 (1992).

18. R.H. Aebersold, Patent WO 9214702, Can, *Compounds and methods for sequencing amino acids* (1992).